FINAL REPORT

TOPICS IN UNCONVENTIONAL IMAGERY

B. Roy Frieden
Tel: (520) 621-4904
Fax: (520) 621-5300
Email: roy.frieden@optics.arizona.edu

SCIENTIFIC PERSONNEL

B.R. Frieden, principal investigator
Sergio Barraza-Felix, Doctoral Research Assistant

THE PROBLEM

The shimmering effects of atmospheric turbulence are familiar to anyone who has looked across the room through the air rising from a hot radiator; or at a distant object across a desert in the summertime. The image appears to be randomly distorted and broken up into characteristic 'speckles'. This is a real effect, not an optical illusion, so that any picture taken through the turbulence with a camera will suffer the same problem. We have studied how to reduce these turbulence effects on such camera images *after they are taken*. A special-purpose digital processing algorithm is used called the 'image division' method.

BASIC ALGORITHM

The processing approach is based upon the use of two *short-exposure* images, i.e. with exposure times the order of .01s or less. Also, the images are taken so far apart in time as to 'see' independent turbulence distributions in the optical pupil of the camera. The two images are Fourier transformed, and these are divided one by the other (thus, 'image division'). This produces the basic data D_n, $n = 1,\ldots,N$ where N is the number of frequencies. The data relate to the two unknown PSFs (point spread functions) $s_n^{(1)}, s_n^{(2)}$ via Eq. (12) of Ref. [1]. (Note: See first appendix at end of report.) These are solved for, along with the two object estimates $o_n^{(1)}, o_n^{(2)}$ via the recursive algorithm shown in Fig. 1 of Ref. [2] (second appendix at end of report).

TESTS OF THE ALGORITHM

The algorithm has been extensively tested on computer simulations [2]. More recently, it has been tested on infrared imagery provided by Kitt Peak National Observatory (K. Hege, J.Garcia). All results are positive. The method works.

We have just obtained stellar speckle imagery from the Univ. of New Mexico, and are about to test the algorithm on this as well.

An interesting result of these tests is that the 'tails' of the image contain significant information about the unknown PSFs. Hence the imagery must not be spatially truncated while still at significant values above background. If these are missing, the algorithm is misled into 'thinking' that there is a missing source located somewhere outside the truncated field. Hence it does not converge.

CONCLUSIONS

The image division method is a working method for effectively and speedily reducing the effects of random turbulence upon given images. All processing may be done on a personal computer (as was the case for the testing above). No large mainframe computer is needed. The algorithm could conceivably be implemented via computer chip in a hand-held 'smart' camera. The camera simply takes two successive images through the turbulence each time the shutter is activated. The two images are then processed by the image division algorithm. An FFT (Fast Fourier transform) chip can be used to implement all needed Fourier transforms, and the output image digitally viewed on a liquid crystal display.

Another possibly useful feature of the approach is an outgrowth of finding the two PSFs. These should give information about the presence of wind shear. High winds ought to cause highly speckled PSFs. Hence the approach could potentially be used to search for wind shear conditions along an imaging path. Since the approach is passive, not requiring active probes such as via lidar or radar, it is simple, cheap and robust. Further testing can ascertain how effective it is versus these active approaches.

Reprinted from

OPTICS COMMUNICATIONS

Optics Communications 150 (1998) 15–21

An exact, linear solution to the problem of imaging through turbulence

B.R. Frieden *

Optical Sciences Center, University of Arizona, Tucson, AZ 85721, USA

Received 10 November 1997; revised 28 January 1998; accepted 28 January 1998

ELSEVIER

Aims and Scope
Optics Communications ensures the rapid publication of contributions in the field of optics and interaction of light with matter.

Abstracted/indexed in:
Chemical Abstracts; Current Contents: Engineering, Computing and Technology; Current Contents: Physical, Chemical & Earth Sciences; EI Compendex Plus; Engineering Index; INSPEC.

Subscription Information 1998
Volumes 145–158 (84 issues) of Optics Communications (ISSN 0030-4018) are scheduled for publication.
Prices are available from the publisher upon request. Subscriptions are accepted on a prepaid basis only. Issues are sent by SAL (Surface Air Lifted) mail wherever this service is available.
For orders, claims, product enquiries (no manuscript enquiries) please contact the Customer Support Department at the Regional Sales Office nearest to you:
New York, Elsevier Science, P.O. Box 945, New York, NY 10159-0945, USA. Tel: (+1) 212-633-3730, [Toll free number for North American customers: 1-888-4ES-INFO (437-4636)], Fax: (+1) 212-633-3680, E-mail: usinfo-f@elsevier.com
Amsterdam, Elsevier Science, P.O. Box 211, 1000 AE Amsterdam, The Netherlands. Tel: (+31) 20-485-3757, Fax: (+31) 20-485-3432, E-mail: nlinfo-f@elsevier.nl
Tokyo, Elsevier Science, 9-15, Higashi-Azabu 1-chome, Minato-ku, Tokyo 106, Japan. Tel: (+81) 3-5561-5033, Fax: (+81) 3-5561-5047, E-mail: info@elsevier.co.jp
Singapore, Elsevier Science, No. 1 Temasek Avenue, #17-01 Millenia Tower, Singapore 039192. Tel: (+65) 434-3727, Fax: (+65) 337-2230, E-mail: asiainfo@elsevier.com.sg
Claims for issues not received should be made within six months of our publication (mailing) date.

US mailing notice – Optics Communications (ISSN 0030-4018) is published semi-monthly by Elsevier Science NL (P.O. Box 211, 1000 AE Amsterdam, The Netherlands). Annual subscription price in the USA is US$ 4325 (valid in North, Central and South America only), including air speed delivery. Periodicals postage paid at Jamaica, NY 11431.
USA Postmasters: Send changes to Optics Communications, Publications Expediting, Inc., 200 Meacham Avenue, Elmont, NY 11003. Airfreight and mailing in the USA by Publications Expediting Inc., 200 Meacham Avenue, Elmont, NY 11003.

The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

Printed in the Netherlands

North-Holland, an imprint of Elsevier Science

ELSEVIER

1 May 1998

ELSEVIER

Optics Communications 150 (1998) 15–21

OPTICS COMMUNICATIONS

An exact, linear solution to the problem of imaging through turbulence

B.R. Frieden *

Optical Sciences Center, University of Arizona, Tucson, AZ 85721, USA

Received 10 November 1997; revised 28 January 1998; accepted 28 January 1998

Abstract

We show how, in principle, to solve the 'blind deconvolution' problem. This is in the context of the problem of imaging through atmospheric turbulence. The approach is digital but not iterative, and requires as input data but two short-exposure intensity images, without the need for reference point sources. By taking the Fourier transform of each image and dividing, a set of linear equations is generated whose unknowns are sampled values of the two random point spread functions that degraded the images. An oversampling by 50% in Fourier space equalizes the number of unknowns and independent equations. With some prior knowledge of spread function support, and in the absence of added noise of image detection, the inverted equations give exact solutions. The two observed images are then inverse filtered to reconstruct the object. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

We first review image sampling theory as it will be used in the processing approach. A knowledge of the fundamentals of incoherent image formation as described, e.g. in Ref. [1], is assumed. For simplicity, we use one-dimensional notation. The theory is trivially extended to two dimensions (e.g. replacing all one-dimensional Fourier sums and integrals with their two-dimensional counterparts). The simulations at the end are two dimensional, as are real images.

2. Sampling theorems

Consider an incoherent object of limited extension $2x_0$. This has an intensity profile $o(x)$, with x the position coordinate, and

$$o(x) = 0 \quad \text{for} |x| \geq x_0. \tag{1}$$

The object is imaged, via a turbulent atmosphere, through a lens system. Although the image $i(x)$ has, theoretically, infinite extension, in practice it has finite extension, since beyond a finite position all intensity values are insignificantly small. For convenience, we place it along the positive x-axis, as shown in Fig. 1. The Fourier transform of the image, the spectrum $I(\omega)$, is sketched on the right (spatial frequency ω is in units of radians/length). For simplicity, only the real part of $I(\omega)$ is shown. Note the cutoff for positive *and negative* frequencies at $\pm\Omega$.

The relation between $i(x)$ and $I(\omega)$ is, of course,

$$I(\omega) = \int_0^\infty \mathrm{d}x\, i(x) \mathrm{e}^{-j\omega x}, \quad j = \sqrt{-1}, \quad -\Omega \leq \omega \leq \Omega. \tag{2}$$

As is well known, the finite cutoff frequency Ω allows us to use, with negligible error, a discrete sum

$$I(\omega) = \sum_{m=0}^{M-1} \Delta x\, i(m\Delta x) \mathrm{e}^{-j\omega m \Delta x}, \quad -\Omega \leq \omega \leq \Omega, \tag{3}$$

where

$$\Delta x = \pi/\Omega, \tag{4}$$

in place of the integral form Eq. (2). Quantity Δx is called the sampling interval, since it is the constant spacing

* E-mail: friedenr@super.arizona.edu

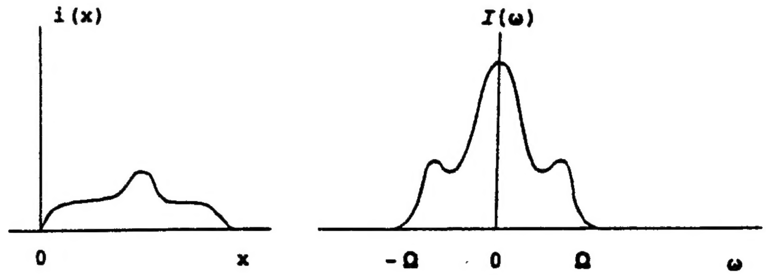

Fig. 1. Image $i(x)$ and its spectrum $I(\omega)$ (real part shown).

between sampled image values $i(m\Delta x), m = 0, 1, \ldots, M - 1$.

Because of the discrete nature of the sum in Eq. (3) the spectrum replicates periodically, $I(\omega + \Omega) = I(\omega)$, as shown in Fig. 2. Note that the interval R contains all of the spectral curve $I(\omega)$, albeit in a rearranged order (i.e. first the positive portion, then the negative portion). This permits us to use, then, as the spectrum

$$I(\omega) = \sum_{m=0}^{M-1} \Delta x i_m e^{-j\omega m \Delta x}, \quad i_m \equiv i(m\Delta x),$$
$$0 \le \omega \le 2\Omega, \tag{5}$$

in place of Eq. (3).

The computed spectrum via Eq. (5) can, in principle, be evaluated at a continuum of frequency values over the given interval. But, for computational purposes, we evaluate it at a fine, discrete subdivision

$$\omega_n \equiv \frac{2n}{N}\Omega, \quad n = 0,1,\ldots,N-1. \tag{6}$$

These discrete values all lie within the interval $(0, 2\Omega)$ prescribed in Eq. (5). The size of N governs the fineness of the frequency sampling. Notice that N can, in fact, be taken as large as is desired. We use this fact in the processing method below.

By Eqs. (4) and (6), the exponent in Eq. (5) becomes

$$j\omega_n m\Delta x = j\frac{2n}{N}\Omega\frac{m\pi}{\Omega} = \frac{2\pi jmn}{N}. \tag{7}$$

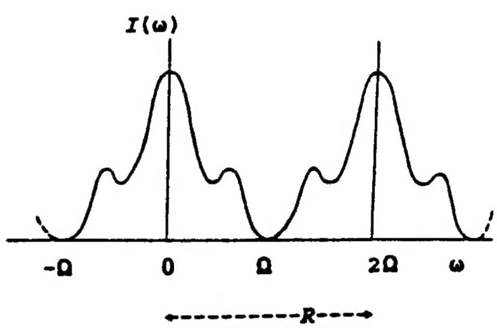

Fig. 2. Replicated image spectrum.

Then the sampled Eq. (5) becomes

$$I(\omega_n) \equiv I_n = \Delta x \sum_{m=0}^{M-1} i_m e^{-2\pi jmn/N}, \quad n = 0,1,\ldots,N-1. \tag{8}$$

This is in the form of a discrete Fourier transform (DFT), but with a difference: there is a generally different number N of output values in frequency space than the M input values in direct space. Our 'division method' of turbulence processing, developed below, hinges upon the use of cases where $N > M$ by appreciable amounts.

3. Image turbulence problem

We now describe, in detail, the image turbulence problem [2–4] at hand. Consider a single object that is imaged twice in succession through random atmospheric turbulence. The images are of short-exposure duration, the order of 1/60 s or less, but are separated in time by more than a short-exposure duration, say 3/60 s. The images, then, 'see' two independent turbulent phase distributions across the optical pupil.

The object has an unknown spectrum $O(\omega_n) \equiv O_n$, $n = 0,1,\ldots,N-1$. Likewise, the two short-exposure images are degraded by two unknown, random optical transfer functions $\tau^{(1)}(\omega_n) \equiv \tau_n^{(1)}$, $\tau^{(2)}(\omega_n) \equiv \tau_n^{(2)}$; $n = 0, 1, \ldots, N - 1$. This might be called a 'triply blind' problem, in view of the three sets of unknowns. The two observed intensity images give rise to two known image spectra $I_n^{(1)}, I_n^{(2)}$, $n = 0,1,\ldots,N-1$ as computed via Eq. (8). The image, object and transfer function spectra are connected by a transfer theorem [1]

$$I_n^{(i)} = \tau_n^{(i)} O_n, \quad i = 1,2; \quad n = 0,1,\ldots,N-1. \tag{9}$$

(For the time being, we ignore added noise of detection.) This allows us to remove one set of unknowns from the problem, as follows. The approach is called the 'image division' method.

4. Image division method

Recall that the image spectra $I^{(1)}, I^{(2)}$ are known as secondary data via Eq. (8) from the primary intensity data $i^{(1)}, i^{(2)}$. Divide the two image spectra, to form new data

$$D_n \equiv \frac{I_n^{(1)}}{I_n^{(2)}} = \frac{\tau_n^{(1)} O_n}{\tau_n^{(2)} O_n} = \frac{\tau_n^{(1)}}{\tau_n^{(2)}}, \quad n = 0,1,\dots,N-1. \qquad (10)$$

The unknowns O_n have dropped out. As in Eq. (8), $\tau^{(1)}$ and $\tau^{(2)}$ relate to unknown short-exposure point spread functions (PSFs) $s^{(1)}$ and $s^{(2)}$ via sampling expressions

$$\tau_n^{(i)} = \Delta x \sum_{m=0}^{M-1} s_m^{(i)} \mathrm{e}^{-2\pi jmn/N}, \quad i = 1,2. \qquad (11)$$

Substituting these into Eq. (10) gives our working expression

$$D_n = \frac{\sum_{m=0}^{M-1} s_m^{(1)} \mathrm{e}^{-2\pi jmn/N}}{\sum_{m=0}^{M-1} s_m^{(2)} \mathrm{e}^{-2\pi jmn/N}}, \quad n = 0,1,\dots,N-1. \qquad (12)$$

The left-hand side is known, as data D_n. The PSFs $s^{(1)}, s^{(2)}$ are the unknowns. Regarded as a computational problem, Eq. (12) represents N equations in $2M$ unknowns. Recalling that N can be made as large as we wish, can it be made large enough compared to $2M$ such that the equations give a solution for the PSFs?

To see what is really going on, cross-multiply Eq. (12) and bring everything over to one side. This gives a set of N equations

$$\sum_{m=0}^{M-1} \left(s_m^{(1)} - D_n s_m^{(2)}\right) \mathrm{e}^{-2\pi jmn/N} = 0, \quad n = 0,1,\dots,N-1. \qquad (13)$$

These are, in fact, *linear* in the unknowns $s^{(1)}, s^{(2)}$. This offers hope for a closed-form, exact solution to the problem. Such a solution is found as follows.

Recall that the data values are the D_n. These are generally complex. Represent them as

$$D_n \equiv M_n \mathrm{e}^{2\pi j\phi_n/N}. \qquad (14)$$

where M_n and ϕ_n are known modulus and phase values. (The factor $2\pi/N$ is there for later convenience.)

Eqs. (13) are complex. The real parts give rise to one set of equations, the imaginary parts to another. Using Eq. (14), these are

$$\text{Re part:} \sum_m s_m^{(1)} \cos\left(\frac{2\pi mn}{N}\right) - \sum_m M_n s_m^{(2)} \cos\left[\frac{2\pi}{N}(\phi_n - mn)\right] = 0,$$

$$\text{Im part:} \sum_m s_m^{(1)} \sin\left(\frac{2\pi mn}{N}\right) + \sum_m M_n s_m^{(2)} \sin\left[\frac{2\pi}{N}(\phi_n - mn)\right] = 0. \qquad (15)$$

These are $2N$ linear equations in the $2M$ unknowns $s^{(1)}, s^{(2)}$.

One hurdle that has to be overcome is that the preceding equations are homogeneous (right-hand sides are 0). Then the direct solution for the case $N = M$ would be that all PSF values are zero. This violates normalization and is not, of course, a physically useful solution. It is better to work with inhomogeneous linear equations. Luckily, the equations may be made inhomogeneous as follows.

The quotient form of the right-hand side of data Eq. (12) shows that the data D_n are only sensitive to the ratios of PSF values. Consequently, knowledge of the D_n only permits the $s^{(1)}, s^{(2)}$ to be known to an arbitrary multiplicative constant. This is acceptable, since if the PSFs are initially off by such a factor they are immediately correctable by normalization. (This models each image, in the usual way, as being a redistribution of the fixed amount of light energy that is present in the object.)

We take advantage of such an arbitrary multiplicative factor in the following way. Any one of the PSF values may be arbitrarily set equal to 1 (say), and the data equations (15) must then give, when solved for the other PSF values, numbers that are correctly ratioed to the 1. For example, if we set $s_{M/2}^{(1)} = 1$ whereas its actual correct value is 0.50 then if the correct value of (say) $s_5^{(2)} = 0.7$ the solution to Eq. (15) would give $s_5^{(2)} = 1.4$. The only problem such a procedure could encounter is to set the value 1 equal to a PSF value which, in reality, is 0. There is no useable multiplier of 0 that brings it up to a finite number. Similarly, if the real PSF value is small, say 0.0001, then setting it equal to 1 would give as solution for the other PSF values very large numbers. Such a solution could be unstable in the presence of small errors in the input image data.

To avoid the latter problems, we equate to 1 a PSF value which is least likely to be 0. Such a one would be at the center of the PSF, i.e. $s_{[M/2]}^{(1)}$ (notation: $[M/2] = M/2$ for M even or $(M-1)/2$ for M odd). The energy should tend to be maximal there. Doing this changes the homogeneous set of equations (15) into the inhomogeneous set

$$\sum_{m \neq [M/2]} s_m^{(1)} \cos\left(\frac{2\pi mn}{N}\right) - \sum_m M_n s_m^{(2)} \cos\left[\frac{2\pi}{N}(\phi_n - mn)\right] = -\cos\left(\frac{2\pi [M/2] n}{N}\right), \quad \text{for } n = 0,1,\dots,N-1,$$

$$\sum_{m \neq [M/2]} s_m^{(1)} \sin\left(\frac{2\pi mn}{N}\right) + \sum_m M_n s_m^{(2)} \sin\left[\frac{2\pi}{N}(\phi_n - mn)\right] = -\sin\left(\frac{2\pi [M/2] n}{N}\right), \quad \text{for } n = 1,2,\dots,N-1. \qquad (16)$$

(The second-set equation for $n = 0$ is the tautology $0 = 0$ and, so, is skipped.)

Before attempting solution, it is necessary to delete all equations which are linear multiples of others. Including them ruins any matrix inversion approach to the problem. It turns out that, if one uses $N = 2M$ equations, then the first $M-1$ equations and the last $M+1$ equations are linearly independent and may be used to form the solution set. We found, by simulation, that the method works. *The solution $s^{(1)}$, $s^{(2)}$ incurs no error.*

The preceding paragraph applies specifically to one-dimensional $(1 \times M)$ images. For the corresponding $M \times M$ problem in two dimensions it was found that an $N \times N$ matrix of frequencies allowed for a unique solution if $N = 1.5M$, i.e. an oversampling by 50% in each direction of frequency space.

5. Demonstrations

To easily judge reconstruction quality, we worked with a simple object – a large figure X of intensity level M – within an $M \times M$ image field of size $M = 16$. To facilitate pixel-by-pixel comparisons of ground truth and reconstructed images, we display each pixel as an easily recognized alphanumeric symbol. These cover the 10 intensity levels (blank).: / % $ Z # @ M, respectively. Small field size M was also dictated by the fact that any solution to the (two-dimensional) problem requires inversion of a matrix of size $2M^2 \times 2M^2$, which rapidly increases with M and is already of size 512×512 for our problem.

The PSFs were randomly generated in accordance with a Kolmogoroff power spectrum [2]. For each problem, a pair of PSFs were so generated, convolved with the X-shaped object to form the two images, and then processed by the image division method. In the first demonstration, no noise of image detection was added. Fig. 3 shows the true PSFs in the left-hand column, and their reconstructions, by the image division method, in the right-hand column. A pixel-by-pixel comparison shows that the reconstructions are nearly exact (to the 10 gray level scale used). Any errors are believed due to round-off during the calculation, which was done in single precision arithmetic.

With the PSFs reconstructed, each was Fourier transformed to give a reconstructed transfer function τ_n. Using Eq. (9), each τ_n was then divided into its corresponding image spectrum to form an estimated object spectrum. This procedure is called 'inverse filtering'. The object spectra were then Fourier transformed back into x-space to give the reconstructed objects. Fig. 4 shows each of the two simulated images in side-by-side comparison with its reconstructed object by the inverse-filtering procedure. It is seen that the reconstructions are much improved over their corresponding images and, in fact, are very good in absolute terms. In an effort to find why the approach works, we empirically varied the support widths of the PSFs. In those

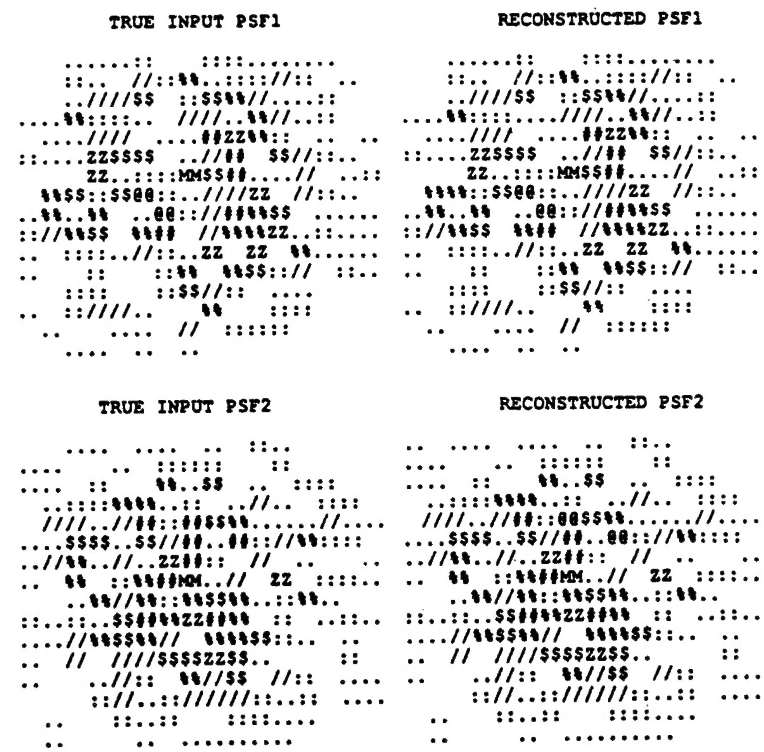

Fig. 3. PSF1, PSF2 and their reconstructions.

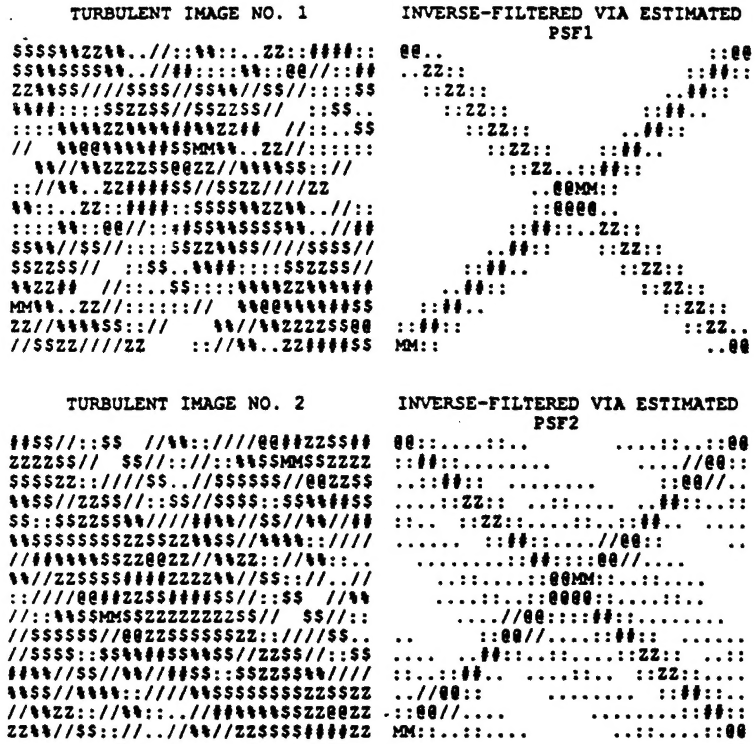

Fig. 4. Turbulent images and their reconstructions (no added noise of detection).

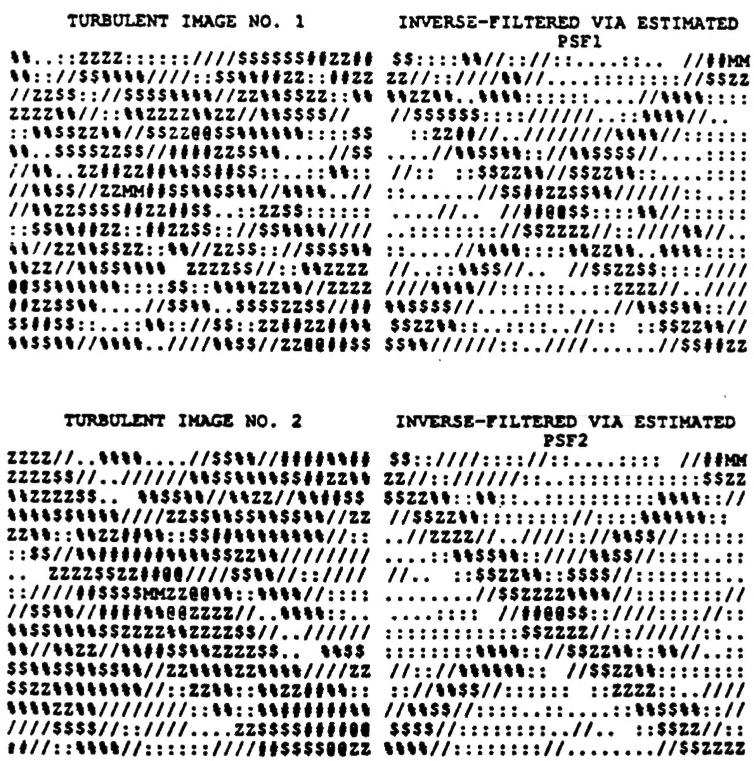

Fig. 5. Turbulent images and their reconstructions (1% additive Gaussian noise of detection).

cases where the two PSFs extend nearly to the edge of the field, it was found that the system of equations (16) has a well-defined inverse, so that the method gives perfect or near-perfect results (as above). Such cases actually represent a situation of strong prior knowledge of the support (region of non-zero values) of the PSFs. That is, the support is known to be approximately the size of the field. Here one can say that the field size is a 'good' (least) upper bound to PSF support.

We found, further, that a range of PSF supports are well-accommodated by the approach. As long as PSF values of about 1% of the maximum or more exist near the edge of the field, then the field size represents strong enough prior knowledge of PSF support to give nearly perfect reconstructions (assuming zero additive noise of detection).

But in cases where the PSF values near the field edge are significantly less than 1% of the maximum value, the PSFs are contracted in support toward the center of the field by significant, and unknown, amounts. Then the known field size no longer represents a 'good' upper bound to the PSF support. In these cases, the resulting system of equations (16) turns out to have a poorly defined inverse; no solution is forthcoming for the PSFs using single-precision arithmetic.

From this we conclude that good prior knowledge of PSF support is important to effective use of the approach. In fact, PSF support can be computed from prior knowledge of the overall strength of the turbulence (as measured by the Fried r_0 extension parameter [5,2], for example). Support knowledge is often a strong aid in the estimation of PSFs. For example, knowledge of the support for the Fourier transform of the PSFs, the OTFs $\tau^{(1)}, \tau^{(2)}$, are a vital part of the blind deconvolution approach of Holmes [6].

In the preceding demonstration, although the two PSFs are random there was no added randomness due to noise of image detection. To test the sensitivity of the approach to such additional noise, we added 1% additive, Gaussian noise to both images and re-processed them. The result was a significant deterioration in the quality of the PSF reconstructions (not shown for brevity), but not enough to wipe out reconstructed object details. Fig. 5 shows the turbulent and (now) noisy images for this case, on the left, along with their two corresponding object reconstructions on the right. The X shape of the object is still clearly visible in both reconstructions.

6. Discussion

As was seen, the image division method, as currently used, works within a domain of 1%, or less, noise of detection. However, as with other poorly posed problems [4], the image division method should be capable of 'regularization', such that it will tolerate noise levels on the order of 10%. Luckily, there are many candidate regularization schemes to choose from [4,7,8]. We believe that this should be the main thrust of future research on the problem.

There may be something to gain in noise de-sensitization by sampling at finer subdivisions than as used in the demonstrations. This would lead to a situation of more equations than unknowns, and the transition to a least-squares solution. Such a solution effects a certain degree of regularization and, so, will be one of the regularization schemes (referred to above) that are experimented with. The main problem that is anticipated with this increased-data approach is the need for higher-rank matrix operations (transposition, multiplication, inversion), which are heavy consumers of computation time. However, there are ways of getting around this problem, such as by posing the inversion problem as a minimization problem and seeking a maximum gradient solution instead. Such an approach would avoid the need for matrix inversion, for example.

The approach seems generalizable to a situation where more than two short-exposure images are at hand. For example, in the case of three images $I^{(1)}, I^{(2)}$ and $I^{(3)}$ two (now) data sets $D_n^{(1)} \equiv I_n^{(1)}/I_n^{(2)}$ and $D_n^{(2)} \equiv I_n^{(2)}/I_n^{(3)}$ (see Eq. (10)) could be formed, the object spectrum O_n would drop out as before, and two sets of linear equations (of the form of Eq. (13)) in, now, the three PSFs $s_m^{(1)}, s_m^{(2)}$ and $s_m^{(3)}$ would result. Combining the two sets of linear equations into one should result in a problem that is soluble by straightforward linear inversion. Given the increased amount of independent data in the problem, a higher degree of regularization should be attainable. The question is whether this benefit would offset the added price to be paid in added stability and tracking needs, as well as image manipulation and storage needs.

This study shows that, both by theory and computer simulation, random atmospheric turbulence is not the ultimate limitor of object resolution quality. It is noise of detection, coupled with lack of knowledge of spread function support.

Acknowledgements

This work was sponsored by a grant from the program on Unconventional Imagery of the Air Force Office of Scientific Research.

References

[1] J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.

[2] F. Roddier, in: E. Wolf (Ed.), Progress in Optics XIX, North-Holland, Amsterdam, 1981.
[3] J.C. Dainty, in: Topics in Applied Physics 9, 2nd ed., Springer, New York, 1984.
[4] J.C. Dainty, J.R. Fienup, in: H. Stark (Ed.), Image Recovery, Academic Press, New York, 1987.
[5] D.L. Fried, J. Opt. Soc. Am. 56 (1966) 1372.
[6] T.J. Holmes, J. Opt. Soc. Am. A 9 (1992) 1052.
[7] P.A. Jansson, Deconvolution of Images and Spectra, 2nd ed., Academic Press, San Diego, 1997.
[8] M.C. Roggemann, B. Welsh, Imaging through Turbulence, CRC, Boca Raton, 1996.

Instructions to Authors (short version)

(A more detailed version of the instructions is published in the preliminary pages of each volume)

Submission of papers

Manuscripts (one original and two copies), should be sent to one of the Editors, whose addresses are given on the inside of the journal cover.

Original material. Submission of a manuscript implies that the paper is not being simultaneously considered for publication elsewhere and that the authors have obtained the authority for publication, if needed.

Refereeing. Submitted papers will be refereed and, if necessary, authors may be invited to revise their manuscript. Authors are encouraged to list the names (addresses and telephone numbers) of up to five individuals outside their institution who are qualified to serve as referees for their paper. The referees selected will not necessarily be from the list suggested by the author.

Types of contributions

The journal Optics Communications publishes short communications and full length articles in the field of optics and quantum electronics.

Short communications are brief reports of significant, original and timely research results that warrant rapid publication. The length of short communications is limited to six journal pages. Proofs will not be mailed to authors prior to publication unless specifically requested.

Full length articles are subject to the same criteria of significance and originality but give a more complete and detailed account of the research results. Proofs of all full length articles will be mailed to the corresponding author, who is requested to return the corrected version to the publisher within two days of receipt.

Manuscript preparation

All manuscripts should be written in good English. The paper copies of the text should be prepared with double line spacing and wide margins, on numbered sheets. See notes opposite on electronic version of manuscripts.

Structure. Please adhere to the following order of presentation: Article title, Author(s), Affiliation(s), Abstract, classification codes (PACS and/or MSC) and keywords, Main text, Acknowledgements, Appendices, References, Figure captions, Tables.

Corresponding author. The name, complete postal address, telephone and Fax numbers and the E-mail address of the corresponding author should be given on the first page of the manuscript.

PACS codes/keywords. Please supply one to four classification codes (PACS and/or MSC) and 1–6 keywords of your own choice for indexing purposes.

References. References to other work should be consecutively numbered in the text using square brackets and listed by number in the Reference list. Please refer to a recent issue of the journal or to the more detailed instructions for examples.

Illustrations

Illustrations should also be submitted in triplicate: one master set and two sets of copies. The *line drawings* in the master set should be original laser printer or plotter output or drawn in black india ink, with careful lettering, large enough (3–5 mm) to remain legible after reduction for printing. The *photographs* should be originals, with somewhat more contrast than is required in the printed version. They should be unmounted unless part of a composite figure. Any scale markers should be inserted on the photograph itself, not drawn below it.

Colour plates. Figures may be published in colour, if this is judged essential by the editor. The publisher and the author will each bear part of the extra costs involved. Further information is available from the full length instructions.

After acceptance

Important. When page proofs are made and sent out to authors, this is in order to check that no undetected errors have arisen in the typesetting (or file conversion) process. No changes in, or additions to, the edited manuscript will be accepted.

Copyright transfer. You will be asked to transfer copyright of the article to the publisher. This transfer will ensure the widest possible dissemination of information.

Electronic manuscripts

The publisher welcomes the receipt of an electronic version of your accepted manuscript (preferably encoded in LaTex). If you have not already supplied the final, revised version of your article (on diskette) to the Journal Editor, you are requested to send a file with the text of the accepted manuscript directly to the Publisher by e-mail or on diskette (allowed formats 3.5" or 5.25" MS-DOS, or 3.5" Macintosh) to the address given below. Please note that no deviations from the version accepted by the Editor of the journal are permissible without the prior and explicit approval by the Editor. Such changes should be clearly indicated on an accompanying printout of the file.

Author benefits

No page charges. Publishing in Optics Communications is free.

Free offprints. The corresponding author will receive 50 offprints free of charge. An offprint order form will be supplied by the publisher for ordering any additional paid offprints.

CONTENTS-Alert. Journal included in this free Elsevier's pre-publication contents alerting service (for information, please contact *C-ALERT.OPTICS@ELSEVIER.NL*).

ContentsDirect. Journal included in this free Elsevier's pre-publication contents alerting service (for information, please visit the Website *http://www.elsevier.nl/locate/ContentsDirect*).

Discount. Contributors to Elsevier Science journals are entitled to a 30% discount on all Elsevier Science books.

Further information (after acceptance)

Elsevier Science B.V., Optics Communications
Issue Management Physics and Astrophysics
P.O. Box 2759, 1000 CT Amsterdam, The Netherlands
Fax: +31 20 4852319
E-mail: PHYSDESK@ELSEVIER.NL

Regularization of the image division approach to blind deconvolution

Sergio Barraza-Felix and B. Roy Frieden

A problem of blind deconvolution arises when one attempts to restore a short-exposure image that has been degraded by random atmospheric turbulence. We attack the problem by using two short-exposure images as data inputs. The Fourier transform of each is taken, and the two are divided. The result is the quotient of the two unknown transfer functions. The latter are expressed, by means of the sampling theorem, as Fourier series in corresponding point-spread functions, the unknowns of the problem. Cross multiplying the division equation gives an equation that is linear in the unknowns. However, the problem has, initially, a multiplicity of solutions. This deficiency is overcome by use of the prior knowledge that the object and the point-spread functions have finite (albeit unknown) support extensions and also are positive. The result is a fixed-length, linear algorithm that is regularized to the presence of 4–15% additive noise of detection. © 1999 Optical Society of America

OCIS codes: 100.3010, 100.5070, 010.1330.

1. Image Sampling

For simplicity the algorithm is first developed for one-dimensional imagery. Following this, the theory is generalized to two dimensions to allow for two-dimensional images.

Consider an incoherent image $i(x)$ that has a sharp cutoff frequency Ω. The finiteness of Ω implies that $i(x)$ has infinite extension. However, for large enough values of $|x|$ it will fall off to negligible values compared with the inevitable noise of detection. For analytical convenience we replace these values with zeros. This is an approximation. Let this truncated image have length L. For convenience, place this image along the positive x axis. Denote the Fourier transform of the original image $i(x)$ by $I_0(\omega)$. The spatial frequency ω is in units of radians/length. Then $I_0(\omega) \sim I(\omega)$, where

$$I(\omega)_{-\Omega\le\omega\le\Omega} = \int_0^L \mathrm{d}x\, i(x)\exp(-j\omega x). \quad (1)$$

The image spectrum can be approximated by the truncated image spectrum.

By the Whittaker–Shannon sampling theorem,[1] the finite cutoff frequency Ω allows us to use in place of Eq. (1), with zero error, the discrete sum

$$I(\omega)_{-\Omega\le\omega\le\Omega} = \sum_{m=0}^{M-1} \Delta x\, i(m\Delta x)\exp(-j\omega m\Delta x). \quad (2)$$

Here $M - 1 = L/\Delta x$ and the fixed interval

$$\Delta x \equiv \pi/\Omega \quad (3)$$

is called the sampling interval, because it spaces the sampled intensity image as values $i(m\Delta x)$.

Because of the discrete nature of the sum in Eq. (2) the spectrum replicates periodically. This permits us to use, then, as the spectrum

$$I(\omega)_{0\le\omega\le2\Omega} = \sum_{m=0}^{M-1} \Delta x\, i_m \exp(-j\omega m\Delta x),$$
$$i_m \equiv i(m\Delta x) \quad (4)$$

in place of Eq. (2).

The spectrum computed with Eq. (4) can, in principle, be evaluated at a continuum of frequency values over the given interval. For computational purposes we evaluate it at a fine, but discrete, subdivision

$$\omega_n \equiv (2n/N)\Omega, \qquad n = 0, 1, \ldots, N-1. \quad (5)$$

The authors are with the Optical Sciences Center, University of Arizona, Tucson, Arizona 85721. B. R. Frieden's e-mail address is friedenr@super.arizona.edu.

Received 2 July 1998; revised manuscript received 19 November 1998.

0003-6935/99/112232-08$15.00/0

These discrete values are all within the interval $(0, 2\Omega)$ prescribed in Eq. (4). The size of N governs the fineness of the frequency sampling. Notice that N can, in fact, be taken as large as is desired. We use this fact in the processing method below.

By Eq. (5), the exponent in Eq. (4) becomes

$$j\omega_n m\Delta x = j\frac{2n}{N}\Omega\frac{m\pi}{\Omega} = \frac{2\pi jmn}{N}. \qquad (6)$$

Then Eq. (4) simplifies to

$$I(\omega_n) \equiv I_n = \Delta x \sum_{m=0}^{M-1} i_m \exp(-2\pi jmn/N),$$
$$n = 0, 1, \ldots, N-1. \qquad (7)$$

Equation (7) is in the form of a discrete Fourier transform but with a difference: generally the number N of output values in frequency space is different from the number M of input values in direct space. Our division method of turbulence processing, developed below, hinges on the use of cases for which $N > M$.

2. Image Turbulence Problem

The problem of imaging through a turbulent atmosphere has a long history.[2-9] The problem is called blind deconvolution because not only is the goal of the deconvolution procedure—the sharp object—unknown but so is the point-spread function (PSF) of the imaging process. A good reference on ordinary (not blind) deconvolution methods is Jansson.[10] Various approaches to solving the blind deconvolution problem have been tried; see, for example, Refs. 5–9. These are generally open-ended iterative searches, growing out of either a gradient search of solution space or a replacement algorithm (successive Fourier transformation and replacement). Also, the approaches are nonlinear in the unknowns (object and PSF) in that these unknowns multiply each other in the imaging equations. Our approach will instead be of fixed length and linear in its unknowns.

The approach grows out of the following imaging situation. A single incoherent object is imaged twice in succession through a turbulent atmosphere. The images are of short-exposure duration, of the order of 1/60 s or less, but are separated in time by more than one short-exposure duration, say, 3/60 s. The images, then, see two fixed, independent turbulent phase distributions across the optical pupil.

The object has an unknown spectrum $O(\omega_n) \equiv O_n$, $n = 0, 1, \ldots, N-1$. Likewise, the two short-exposure images are degraded by two unknown, random optical transfer functions $\tau^{(1)}(\omega_n) \equiv \tau_n^{(1)}$, $\tau^{(2)}(\omega_n) \equiv \tau_n^{(2)}$; $n = 0, 1, \ldots, N-1$. This might be called a triply blind problem in view of the three sets of unknowns. The two observed intensity images give rise to two known image spectra, $I_n^{(1)}$ and $I_n^{(2)}$, $n = 0, 1, \ldots, N-1$, as computed from Eq. (7). As is known,[1] the image, object, and transfer function spectra are connected by a transfer theorem,

$$I_n^{(i)} = \tau_n^{(i)} O_n, \qquad i = 1, 2; \quad n = 0, 1, \ldots, N-1. \qquad (8)$$

(For the time being, we ignore added noise of detection.) This allows us to remove one set of unknowns from the problem, as follows.

3. Image Division Method

For convenience, we use the shorthand notation that $I^{(1)}$ means $I_n^{(1)}$, $n = 0, 1, \ldots, N-1$, and similarly for $I^{(2)}$. Recall that the image spectra $I^{(1)}$ and $I^{(2)}$ are known as secondary data [from Eq. (7)] from the primary intensity data $i^{(1)}$ and $i^{(2)}$. Form the quotient of the secondary data:

$$D_n \equiv \frac{I_n^{(1)}}{I_n^{(2)}} = \frac{\tau_n^{(1)} O_n}{\tau_n^{(2)} O_n} = \frac{\tau_n^{(1)}}{\tau_n^{(2)}}, \qquad n = 0, 1, \ldots, N-1. \quad (9)$$

(Thus, the "image division" method.) Equation (8) was used. Note that the object spectrum O_n canceled out in Eq. (9), which eliminates one set of unknowns from the problem. Such cancellation hinges on a lack of image detector noise in Eq. (8).

Now, as at Eq. (7), $\tau^{(1)}$ and $\tau^{(2)}$ relate to unknown short-exposure PSF's $s^{(1)}$ and $s^{(2)}$ by means of the sampling expression

$$\tau_n^{(i)} = \Delta x \sum_{m=0}^{M-1} s_m^{(i)} \exp(-2\pi jmn/N), \qquad i = 1, 2. \quad (10)$$

Substituting Eq. (10) into Eq. (9) gives our working expression:

$$D_n = \sum_{m=0}^{M-1} s_m^{(1)} \exp(-2\pi jmn/N) \Big/ \sum_{m=0}^{M-1} s_m^{(2)} \exp(-2\pi jmn/N),$$
$$n = 0, 1, \ldots, N-1. \qquad (11)$$

The left-hand side of Eq. (11) is known as data. The PSF's $s^{(1)}$ and $s^{(2)}$ are the unknowns. Regarded as a computational problem, Eq. (11) represents N equations in $2M$ unknowns. Recalling that N can be made as large as we wish, can it be made large enough compared with $2M$ that the equations give a solution for the PSF's?

To see what is really going on, cross multiply Eq. (11) and bring everything over to one side, which yields a set of N equations

$$\sum_{m=0}^{M-1} [s_m^{(1)} - D_n s_m^{(2)}] \exp(-2\pi jmn/N) = 0,$$
$$n = 0, 1, \ldots, N-1. \qquad (12)$$

These are, in fact, linear in the unknowns $s^{(1)}$ and $s^{(2)}$ and thus offer hope for a closed-form solution to the problem.

Recall that the data values are the D_n. These are generally complex. Represent them as

$$D_n \equiv M_n \exp(2\pi j\phi_n/N), \qquad (13)$$

where M_n and ϕ_n are known modulus and phase values. (The factor $2\pi/N$ is there for later convenience.)

Equation (12) is complex. The real parts of Eq.

(12) give rise to one set of equations; the imaginary parts, to another. By Eq. (13), these are

$$\text{Re part:}\quad \sum_m s_m^{(1)} \cos\left(\frac{2\pi mn}{N}\right) - \sum_m M_n s_m^{(2)} \times \cos\left[\frac{2\pi}{N}(\phi_n - mn)\right] = 0,$$

$$\text{Im part:}\quad \sum_m s_m^{(1)} \sin\left(\frac{2\pi mn}{N}\right) + \sum_m M_n s_m^{(2)} \times \sin\left[\frac{2\pi}{N}(\phi_n - mn)\right] = 0. \quad (14)$$

These are $2N$ linear equations in the $2M$ unknowns $s^{(1)}$ and $s^{(2)}$.

4. Overcoming the Homogeneity Problem

One hurdle that has to be overcome is that the preceding equations are homogeneous (the right-hand sides are 0). We need a right-hand-side vector of known values if the system of equations is to have a solution by conventional linear inversion. Another problem is one of uniqueness of solution: Not only do the true PSF's $s^{(1)}$ and $s^{(2)}$ satisfy the equations but so do the convolutions $s^{(1)}(x) \otimes f(x)$ and $s^{(2)}(x) \otimes f(x)$ with any kernel function $f(x)$. [By Eq. (9), the noise-free data D_n are invariant to multiplication of each of $\tau_n^{(1)}$ and $\tau_n^{(2)}$ by an arbitrary function F_n in frequency space.] The implication is that the matrix of coefficients in Eqs. (14) is rank deficient if the data M_n, ϕ_n contain no noise. This point is clarified in Section 5.

This solution-redundancy problem has to be overcome because the plan is, by Eq. (8), to inverse filter each of the images $I_n^{(1)}$ and $I_n^{(2)}$ with their estimated transfer functions $\tau_n^{(1)}$ and $\tau_n^{(2)}$, respectively. If each of the latter functions is incorrect by an arbitrary factor F_n, this will spoil the inverse-filtered outputs. In fact, this problem will be overcome in the processing algorithm defined in Section 6 below.

The inhomogeneity problem can be overcome as follows: Aside from being invariant to the operation of convolution (as mentioned above), the noise-free PSF solutions are invariant to simple multiplication by a constant [again, see Eq. (9)]. This is acceptable, because if the PSF's are initially incorrect by such a factor then subsequent inverse filtering that uses them will also be merely incorrect by a constant factor. This discrepancy is acceptable because images are adjusted in absolute brightness for the viewer's convenience anyhow.

We take advantage of such an arbitrary multiplicative factor in the following way: The total energies in $s^{(1)}$ and $s^{(2)}$ are

$$E^{(i)} \equiv \sum_m s_m^{(i)}, \qquad i = 1, 2. \quad (15)$$

We may assume any convenient value for $E^{(1)}$. Doing so merely scales $s^{(1)}$ by an arbitrary factor—acceptable by the preceding argument. Then Eq. (15) is solved for any one $s^{(1)}$ value, say,

$$s_k^{(1)} = E^{(1)} - \sum_{m \neq k} s_m^{(1)}. \quad (16)$$

This $s^{(1)}$ value is removed from the left-hand sums in Eqs. (14), because it is no longer an unknown. Instead, it forms a nonzero right-hand side. Accordingly, Eqs. (14) become

$$\sum_{m\neq k} s_m^{(1)}\left[\cos\left(\frac{2\pi mn}{N}\right) - \cos\left(\frac{2\pi kn}{N}\right)\right] - \sum_m M_m s_m^{(2)} \times \cos\frac{2\pi}{N}(\phi_n - mn) = -E^{(1)}\cos\left(\frac{2\pi kn}{N}\right),$$

$$\sum_{m\neq k} s_m^{(1)}\left[\sin\left(\frac{2\pi mn}{N}\right) - \sin\left(\frac{2\pi kn}{N}\right)\right] + \sum_m M_m s_m^{(2)} \times \sin\frac{2\pi}{N}(\phi_n - mn) = -E^{(1)}\sin\left(\frac{2\pi kn}{N}\right). \quad (17)$$

Index $n = 0, 1, \ldots, N-1$ in the first of Eqs. (17) and $n = 1, 2, \ldots, N-1$ in the second (for $n = 0$ it becomes the tautology $0 = 0$ and, so, is skipped). The equations are no longer homogeneous, as required. They are also no longer rank deficient for M small enough (see Sections 5 and 7).

5. Least-Squares Solution

Equations (17) are linear in the unknowns $s_m^{(1)}$ and $s_m^{(2)}$. They take the convenient matrix form

$$[H]\mathbf{x} = \mathbf{b}, \quad (18)$$

where $\mathbf{x}$ designates the vector of unknowns $s_m^{(1)}$ and $s_m^{(2)}$, $[H]$ denotes the matrix of trigonometric coefficients in Eqs. (17), and $\mathbf{b}$ denotes the vector right-hand sides of Eqs. (17).

Recall that the number N of data frequencies can be made arbitrarily large. With a fixed number $2M - 1$ of unknowns, the problem [Eq. (18)] becomes overdetermined, which allows a least-squares solution to be sought. A least-squares solution is beneficial because it effects a degree of data noise smoothing, i.e., regularization, on its own. However, we also noted that the matrix of coefficients in Eqs. (14) is rank deficient. This deficiency can be overcome by the proper choice of dimension M in Eqs. (17), as described next.

It was previously found[11] that Eqs. (18) can be solved if the user has, effectively, prior knowledge of support regions for the PSF's. Then a least-squares solution to the problem of Eq. (18) can be formed. As much as approximately 2% additive Gaussian noise can be so tolerated.[11] However, this tolerance is too narrow for practical problems. Something else needs to be inserted into the algorithm, namely, the use of prior knowledge[10] about the PSF's, to regularize the algorithm to the presence of greater amounts of noise.

Generally speaking, the more prior knowledge there is about the unknowns of a problem, the better

is the degree of regularization that can be attained.[10] Two forms of prior knowledge are at hand here.[8] One is finite support regions for the PSF's, as discussed above. Another is positivity. Because our images are incoherent, the object and the PSF's represent energy distributions and, hence, must obey a condition of positivity:

$$o_m \geq 0, \qquad s_m^{(1)} \geq 0, \qquad s_m^{(2)} \geq 0. \tag{19}$$

We return to the knowledge of the finiteness of support. Both the object and the PSF's are, effectively, zero outside regions of finite extension. Also, consider only a case in which these extensions are small enough that the image field contains all the image energy: None of it spills outside the field.

Generally denote the true value of a support extension by K and an estimate by $\hat{K}$. The object is assumed, for simplicity, to lie within a square field of linear extension K_{ob}. The two PSF's have, in general, different x-component supports and different y-component supports. However, by a trick [see Eqs. (27) below] we can make them both have effectively the same x-component support, denoted K_x, and the same y-component support, denoted K_y.

This knowledge of prior support mathematically affects the rank of problem (18), as follows: Suppose that both PSF's $s_m^{(1)}$ and $s_m^{(2)}$ have a true common support value K. Also, let the estimated common support value be $\hat{K}$. Then the upper value for m in the sums in Eqs. (17) will now be $\hat{K}$ instead of the image support value M. This means that, effectively, many of the PSF values are equated to zero and, consequently, the matrix $[H]$ becomes narrower.

Matrix $[H]$ is full rank or rank deficient, depending on the presumed support size $\hat{K}$, as follows: First consider a case of noiseless data. If the guess is that $\hat{K} = M$, then any PSF answers are permitted that are wider than true support value K through convolutions with any kernel function $f(x)$ (see the beginning of Section 4) of support length $(M - K)$. Matrix $[H]$ mirrors this ambiguous situation by not permitting a unique solution, i.e., by being rank deficient. Likewise, if $\hat{K}$ is made any value larger than the true support value K, the matrix will still be rank deficient. [Now the kernel function need have only support length $(\hat{K} - K)$.] However, once $\hat{K} \leq K$ the preceding convolution effect can no longer hold (the kernel would necessarily have negative support). These arguments were borne out in computer simulations as described in Section 7.

On the other hand, with data noise present there is no problem of rank deficiency for $[H]$. Rank deficiency followed ultimately from the fact that the noise-free data D_n defined by Eq. (9) are invariant to multiplication of both transfer functions $\tau_n^{(1)}$ and $\tau_n^{(2)}$ by an arbitrary function. However, with data noise present in the image, the division of their spectra $I_n^{(1)}$ and $I_n^{(2)}$ in Eq. (9) no longer results in a simple quotient of transfer functions $\tau_n^{(1)}/\tau_n^{(2)}$. Hence there is no longer a possible ambiguity in the data owing to multiplication of the transfer functions by an arbitrary function. In essence, the noise breaks the ambiguity. Correspondingly, matrix $[H]$ should be full rank for any choice of support $\hat{K}$ as long as the latter does not exceed the image support size M. Again, this behavior was followed in the computer simulations.

With presumed support values $\hat{K}_x$ and $\hat{K}_y$, problem (18) becomes

$$[H_{\hat{K}}]\mathbf{x}_{\hat{K}} = \mathbf{b}_{\hat{K}}. \tag{20}$$

Equation (20) can be solved in the least-squares sense,

$$\|[H_{\hat{K}}]\mathbf{x}_{\hat{K}} - \mathbf{b}_{\hat{K}}\| = \text{min.}, \tag{21}$$

as follows:

We used one-dimensional notation above. Now we proceed to the full, two-dimensional case. It is easy to see that equations of the same form as Eq. (20) result, provided that the now doubly subscripted unknowns $s_{mn}^{(1)}$ and $s_{mn}^{(2)}$ are packed as a new, one-dimensional vector $\mathbf{x}_{\hat{K}}$ of length $M = \hat{K}_x\hat{K}_y$. The number of unknowns is. as in one dimension, $(2M - 1)$. Also, N is now the chosen number of two-dimensional frequencies (ω_m, ω_n). As in one dimension there are $2N - 1$ equations. However, $(N - 1)$ of these equations are found to be repetitions and so are ignored. This leaves, then, N equations in the $(2M - 1)$ unknowns. However, if the support lengths $\hat{K}_x$ and $\hat{K}_y$ are chosen small enough relative to N, the net matrix $[H_{\hat{K}}]$ will be tall and full rank, so a least-squares solution to Eq. (21) can be sought.

Examples of least-squares solutions are provided by the simulations below. There the image spectral field is of dimension $N = 24 \times 24 = 576$ frequencies. Thus there are 576 equations in problem (20). Also, the spatial image field is made to be 16×16 pixels. Hence any PSF with a support length of 16 or more will cause aliasing effects. The true PSF supports are made to be $K_x = K_y = 8$ pixels. The status of the matrix $[H_{\hat{K}}]$ with respect to rank sufficiency is found to be as follows: When no noise is added to the image data the matrix is full rank when $\hat{K}_x \leq 8$ and $\hat{K}_y \leq 8$. With added noise, the matrix is full rank when both $\hat{K}_x \leq 15$ and $\hat{K}_y \leq 15$. The noise breaks the ambiguity problem, as was mentioned above.

6. Net Algorithm

The preceding considerations provide the basis for a fixed-length restoration approach. The approach would be simply the preceding least-squares method if the correct supports $\hat{K}_x$ and $\hat{K}_y$ were known. However, of course they are not. Therefore the plan is to form a series of least-square solutions $\mathbf{x}_{LSQ}$ to Eq. (21), using a fixed sequence of trial values $\hat{K}_x$ and $\hat{K}_y$, and then in some way to judge which solution is the best of the lot. The best solution is defined to be the one that gives a pair of image estimates that best agree with the two given images, subject to enforced positivity [inequalities (19)] on the running object and PSF estimates. The algorithm is given next and, for the reader's convenience, is also shown schematically in Fig. 1:

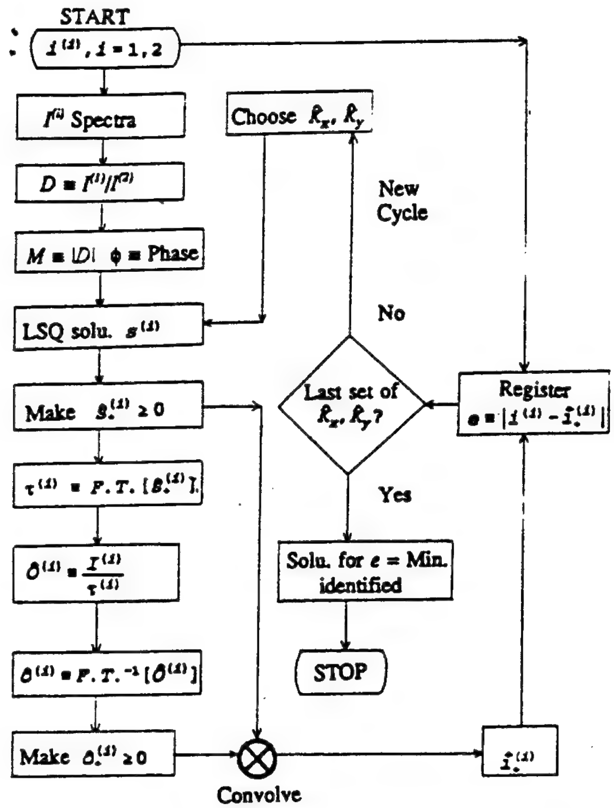

Fig. 1. Flow chart for the image division program.

(1) A pair of short-exposure images $i^{(1)}$ and $i^{(2)}$ is given. These are the primary data for the problem. From these images, spectra $I^{(1)}$ and $I^{(2)}$ are formed by a discrete Fourier transform, Eq. (7). These spectra are divided, as in Eq. (9), to form quantities D_n. From Eq. (13), the associated modulus and phase quantities M_n, and ϕ_n are formed. The latter may be regarded as secondary (i.e., preprocessed) data for the problem.

(2) A pair of support estimates $\hat{K}_x$ and $\hat{K}_y$ is chosen. These supports are both, of course, smaller than the total image field extension, so this choice reduces the number of unknowns $\mathbf{x}$ in problem (21), as discussed above. The support choice strongly affects the solution.

(3) The least-squares solution $\mathbf{x}_{\text{LSQ}} \equiv \hat{s}^{(1)}, \hat{s}^{(2)}$ to problem (21) is formed by use of orthogonal-triangular factorization.[12] This factorization avoids the need to take matrix inverses, a numerically unstable operation for poorly conditioned matrices. The PSF outputs are corrected to obey positivity (19). The rule is to replace all negative values by 0. Call the outputs $\hat{s}_+^{(1)}$ and $\hat{s}_+^{(2)}$.

(4) The discrete Fourier transforms [Eq. (7)] of these outputs, $\tau^{(1)}$ and $\tau^{(2)}$, are taken and are used in Eq. (8) to form object estimates:

$$\hat{O}^{(i)} \equiv I^{(i)}/\tau^{(i)}, \qquad i = 1, 2 \tag{22}$$

by inverse filtering.

(5) These estimates are inverse-Fourier transformed into (x, y) space to yield object estimates $\hat{o}^{(i)}$, $i = 1, 2$.

(6) These object estimates, in general, will not obey positivity (19). We enforce positivity, as in step (3), i.e., by simply zeroing every negative value.

(7) The object should have a support that is consistent with that of the image and that of the PSF's. Define (for example) the x-component support $K_{\text{im}x}$ in the image as its linear extension at the 2% level of maximum intensity. Then, with the trial value $\hat{K}_x$ serving as the x-component support of the PSF's, the x-component support value $\hat{K}_{\text{ob}x}$ of the object is taken to obey

$$\hat{K}_{\text{ob}x} \equiv K_{\text{im}x} - \hat{K}_x + 1. \tag{23}$$

The estimated object is zeroed outside this support interval. The same operations are performed for the y-component direction.

The outputs of steps (6) and (7) are called $\hat{o}_+^{(i)}$, $i = 1, 2$.

(8) Each estimated PSF from step (3) is convolved with a corresponding object estimate from step (7), forming estimated images

$$\hat{i}_+^{(i)} \equiv \hat{s}_+^{(i)} \otimes \hat{o}_+^{(i)}, \qquad i = 1, 2. \tag{24}$$

The symbol $\otimes$ denotes a convolution operation.

(9) The extent to which these estimated images agree with the given images $i^{(1)}$ and $i^{(2)}$ defines how valid were the hypothetical support values in step (2). For example, poor agreement is taken to imply invalid support values. We quantify poor agreement by forming an error metric over both images:

$$e \equiv e^{(1)} + e^{(2)}, \qquad e^{(i)} \equiv \frac{|i^{(i)} - \hat{i}_+^{(i)}|}{|i^{(i)}|}. \tag{25}$$

The vertical lines signify the absolute value operation on the vector within. Parameter e measures the inconsistency of the solution with the image data.

(10) The value of e is registered, and a new cycle begins at step (2), with a new choice of supports.

(11) The minimum value of e that is found over the range of test supports is taken to identify the final solution $\hat{s}_+^{(i)}, \hat{o}_+^{(i)}, i = 1, 2$. The arithmetic average of the two object estimates defines the output object:

$$\hat{o} \equiv ½(\hat{o}_+^{(1)} + \hat{o}_+^{(2)}). \tag{26}$$

7. Demonstrations

The effectiveness of the image division algorithm [steps (1)–(11)] is tested by computer simulation. The computer is an Acer Model 2016 Emerald Desktop Computer. It has a 166-MHz Pentium processor and a 16-MB EDO memory. Execution time is ~1 min for each cycle [steps (1)–(11)] of the algorithm. For the 16 × 16 pixel field used, a typical output requires 120 cycles, or 120 min in all.

The image spectral field size in all cases is 24×24 frequencies, or $N = 576$. (Larger field sizes might require the use of a mainframe computer.) The subdivision in image space is 16×16 pixels, and we confine the object to an 8×8 pixel square region in the center of the field. The two PSF's are constructed with support values $K_x = K_y = 8$.

In the solution search loop we allow for all possible support pairs $(\hat{K}_x, \hat{K}_y)$ over the full range of values $2 \leq \hat{K}_x \leq 15$ and $2 \leq \hat{K}_y \leq 15$ when the simulated data contain noise, or values $\hat{K}_x \leq 8$ and $\hat{K}_y \leq 8$ when there is no noise. For these ranges of support values, matrix $[H_{\hat{K}}]$ is found to be of full rank (as discussed in Section 5). In actual practice, when it is not known whether the image data contain significant noise the full range of supports would be used, of course. A state of rank deficiency is made known to the user by means of a flag indicated by the MATLAB system in use, so this trial support combination is simply skipped in the search procedure.

Additive noise of detection is included in the image simulations. Hence there are two sources of randomness to overcome: the randomness of the PSF's and of the detector noise. The latter is independent Gaussian, with a standard deviation σ that is expressed as a percent of the root-mean-square signal level. For example, 4% noise means that σ is 4% of the root-mean-square image level.

In the first tests, a letter C object is used as the ideal input; see Fig. 2(a). Note the gradual brightening from top to bottom and the sharp edges bounding the letter. These features allow the algorithm to be tested against both subtly changing and rapidly changing gray levels.

The two PSF's for the 4% noise case are shown in Figs. 3(a) and 3(b). These PSF's were constructed independently at each pixel within a central 8×8 field. Hence the true PSF support values (K_x, K_y) are (8, 8). Each pixel intensity is randomly chosen from a uniform probability law, with thresholding at a finite lower value to make the black background intrude randomly within the support boundary. The aim is to make the boundary less well-defined as a square (8×8) entity, so the support information (8, 8) does not represent an unrealistically strong amount of information. The PSF's shown are typical of the ones tested. Since the PSF's were generated independently at each pixel, their power spectra are flat. Real turbulence is, of course, approximated by a Kolmogorov power spectrum.[2] We judged that a flat power spectrum would be a harsher test of the approach than the Kolmogorov spectrum because uncorrelated PSF signals represent a situation of maximum ignorance regarding PSF values. Also, uncorrelated PSF's were easier to generate by computer. Future tests of the algorithm will be made with real (turbulence-degraded) data, which is of course the truest test of any algorithm.

Reconstructions of the letter C by the use of the algorithm are given in Figs. 2(b)–2(d). These reconstructions are for various levels of image noise, as indicated. It is seen that the edge gradients are well

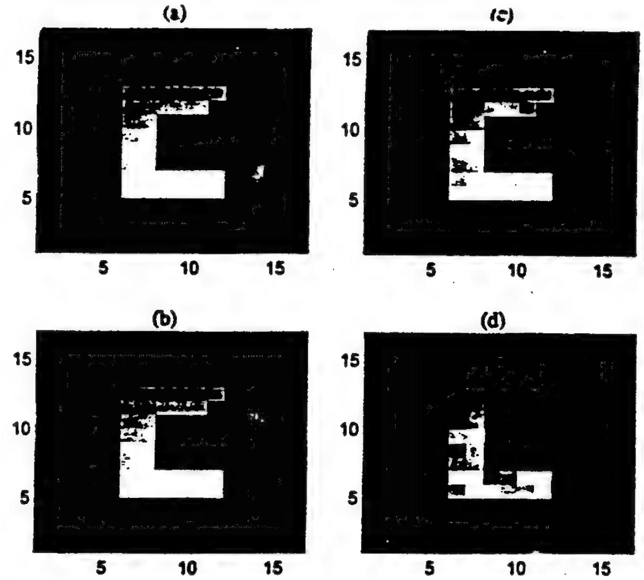

Fig. 2. Object reconstructions at given levels of detector noise (the indicated coordinates are pixels): (a) letter C object, (b) reconstruction with 0% detection noise, (c) reconstruction with 2% noise, (d) reconstruction with 4% noise.

restored at these noise levels, although the subtle gray-level transition is poorly reconstructed at the 4% level of noise.

The reconstructed PSF's are shown, respectively, in Figs. 3(c) and 3(d). It is seen that at this level of noise there is a tendency, although not a strong one, for the pixel values that are internal to the PSF's to be faithfully reconstructed. On the other hand, the supports for the PSF's in Figs. 3(a) and 3(b) are perfectly estimated in the reconstructions in Figs. 3(c) and 3(d). Probably the fidelity of the reconstructions owes more to the correct estimation of the sup-

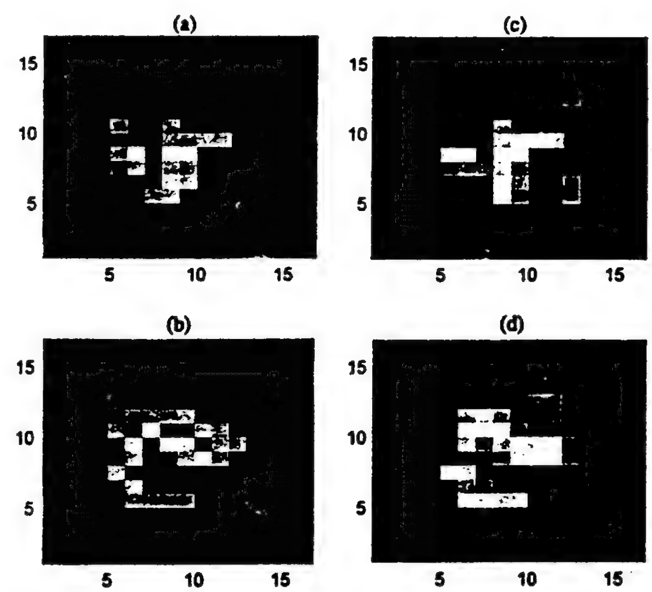

Fig. 3. Data PSF's and their reconstructions for 4% noise: (a) PSF 1, (b) PSF 2, (c) reconstruction of PSF 1, (d) reconstruction of PSF 2.

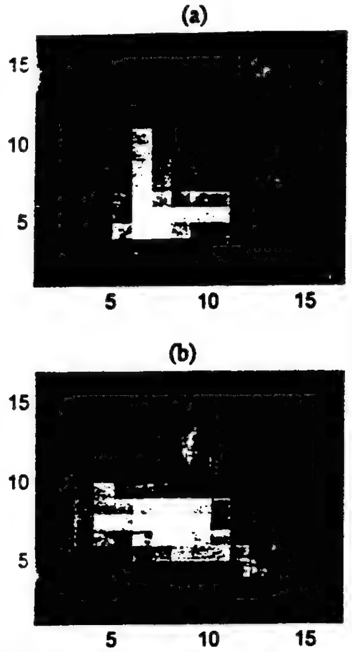

Fig. 4. Data images for 4% noise: (a) image 1, (b) image 2.

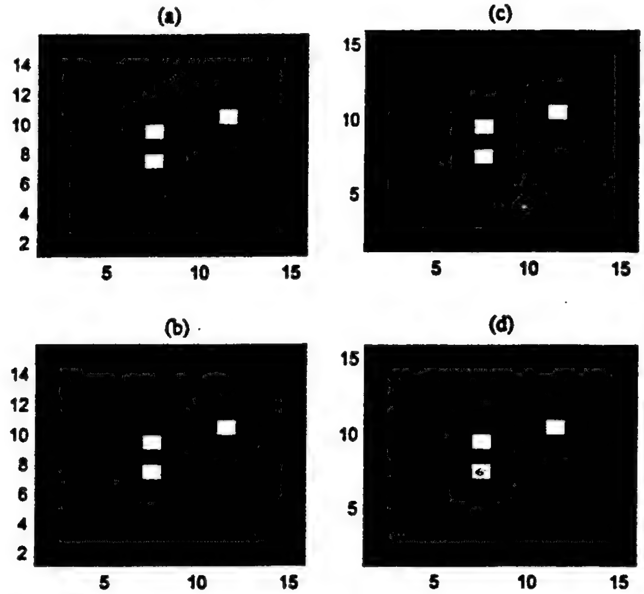

Fig. 5. Impulse object case studies: (a) object, (b) reconstruction with 5% detection noise, (c) reconstruction with 10% noise, (d) reconstruction with 15% noise.

ports than to reproduction of the internal features of the PSF's. Again, we tried to minimize this effect by allowing the zero background to intrude randomly within the 8 × 8 fields of the simulated PSF's.

The two images that were used as inputs in the 4% noise case are shown in Fig. 4. These are, mathematically, the convolutions of the letter C of Fig. 2(a) with the PSF's of Figs. 3(a) and 3(b), plus 4% noise. The images exhibit little resemblance to the letter C and so constitute a good test case for the algorithm.

It is by now well known that pointlike objects are restored exceptionally well when they are constrained to obey positivity.[10,13] We therefore also tested the algorithm against a cluster of point-source objects, which are shown in Fig. 5(a). The image-division algorithm was applied to image pairs of this object at three noise levels: 5%, 10%, and 15%. The object reconstructions are shown, respectively, for these noise cases in Figs. 5(b), 5(c), and 5(d). Indeed, these figures show the three point sources quite well, i.e., with almost no blur. As a negative aspect, spurious background details emerge. These would probably be minimized by the use of a nonlinear restoring technique.

A premise of the algorithm [see step (11)] is that the data inconsistency e has a grand minimum at the correct support levels (K_x, K_y) for the case. To check out this assumption we have plotted, in Fig. 6, e versus $\hat{K}_x$ for various values of $\hat{K}_y$ for 4% noise. From Figs. 3(a) and 3(b) the true support values are here $(K_x, K_y) = (8, 8)$. It is seen that the minimum value is at the low point of the open-circle curve, i.e., for the point $(\hat{K}_x, \hat{K}_y) = (8, 8)$. These are the correct supports for this problem. Notice, however, that the asterisk curve (for $\hat{K}_y = 7$) has a minimum that is nearly as low as that of the open-circle curve. This problem worsens at higher-noise cases, for which the minimum may no longer define the true supports.

8. Discussion

The two PSF's might not have the same support K_x and the same support K_y. In this general case there are four support parameters to vary in the algorithm loop (1)–(11). This case represents a four-dimensional space of solutions, computationally a much more difficult problem to cope with than the two-dimensional one that we used. In fact, we can convert the four-dimensional problem into a two-dimensional one, as follows. Choose weights a_1, a_2, b_1, $b_2 > 0$ such that new images are formed:

$$\tilde{i}^{(1)} \equiv a_1 i^{(1)} + a_2 i^{(2)},$$
$$\tilde{i}^{(2)} \equiv b_1 i^{(1)} + b_2 i^{(2)}, \qquad a_1 + a_2 = 1, \quad b_1 + b_2 = 1. \tag{27}$$

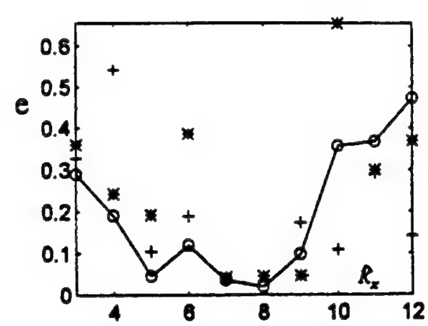

Fig. 6. Does data inconsistency e attain a grand minimum at the correct support levels? Inconsistency e is plotted versus support component $\hat{K}_x$ for various values of component $\hat{K}_y$: $\hat{K}_y = 9$ (pluses), $\hat{K}_y = 8$ (open circles), $\hat{K}_y = 7$ (asterisks). The correct support levels $(\hat{K}_x, \hat{K}_y) = (8, 8)$ are indeed attained at the grand minimum (see the open-circle curve).

The new images $\tilde{i}^{(1)}$ and $\tilde{i}^{(2)}$ are used as the inputs to the image division algorithm. It is apparent from Eqs. (27) that the new images are formed by net PSF's that obey, respectively,

$$\tilde{s}^{(1)} = a_1 s^{(1)} + a_2 s^{(2)},$$
$$\tilde{s}^{(2)} = b_1 s^{(1)} + b_2 s^{(2)}. \qquad (28)$$

Because each of the new PSF's is a weighted sum of both old PSF's, the new PSF's must both have the same x-component support and the same y-component support, as we wanted.

The overall objective of regularizing the image division method has been partially achieved. Depending on the type of object present, anywhere from 4% to 15% noise can be tolerated by the amended approach. The advantages of the approach are the fidelity of its outputs, its linearity and hence potential speed, that it is of fixed length and avoids an open-ended search of solution space, and that it needs but two short-exposure images as inputs. The central role played by the existence of finite supports for the PSF's and the object has become apparent. It results that any prior knowledge of PSF support that can be built into the algorithm to narrow its support search will increase its utility. Also, the method by which positivity is enforced can be improved. The zero-replacement approach of steps (3) and (6) of the algorithm does not permit data consistency in the estimates. Recourse to a nonlinear approach such as maximum entropy[10,13] permits data consistency. This should lead, ultimately, to a better estimate of the PSF supports and, hence, to a better output reconstruction $\hat{o}$.

It has been noted that a practical technique for dealing with the levels of noise encountered in real data has yet to be found.[9] The image division algorithm might achieve this overall aim. The next tests of the algorithm will be on real atmospherically degraded images.

This study was supported by a grant from the Unconventional Imagery program of the U.S. Air Force Office of Scientific Research.

References

1. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
2. F. Roddier, "The effects of atmospheric turbulence in optical astronomy," in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1981), Vol. XIX, pp. 281–376.
3. J. C. Dainty, "Stellar speckle interferometry," in *Laser Speckle and Related Phenomena*, 2nd ed., J. C. Dainty, ed., Vol. 9 of Topics in Applied Physics (Springer-Verlag, New York, 1984), pp. 255–320.
4. J. C. Dainty and J. R. Fienup, "Phase retrieval and image reconstruction for astronomy," in *Image Recovery, Theory and Application*, H. Stark, ed. (Academic, New York, 1987), pp. 231–275.
5. R. S. Lawrence and J. W. Strohbehn, "A survey of clear-air propagation effects relevant to optical communications," Proc. IEEE **58**, 1523–1544 (1970).
6. T. J. Holmes, "Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach," J. Opt. Soc. Am. A **9**, 1052–1061 (1992).
7. M. C. Roggemann and B. Welsh, *Imaging through Turbulence* (CRC Press, Boca Raton, Fla., 1996).
8. G. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," Opt. Lett. **13**, 547–549 (1988).
9. R. G. Lane, "Blind deconvolution of speckle images," J. Opt. Soc. Am. A **9**, 1508–1524 (1992).
10. P. A. Jansson, *Deconvolution of Images and Spectra*, 2nd ed. (Academic, San Diego, Calif., 1997).
11. B. R. Frieden, "An exact, linear solution to the problem of imaging through turbulence," Opt. Commun. **150**, 15–21 (1998).
12. D. Fowley, M. Horton, and J. Scordato, eds., MATLAB, Student Edition, Version 4 (Prentice-Hall, Englewood Cliffs, N. J. 1995).
13. B. R. Frieden, "Image enhancement and restoration," in *Picture Processing and Digital Filtering*, T. S. Huang, ed., Vol. 6 of Topics in Applied Physics (Springer-Verlag, New York, 1975), p. 221.

Foundations of Physics, Vol. 29, No. 10, 1999

F-Information, a Unitless Variant of Fisher Information

B. Roy Frieden[1]

Received January 15, 1999; revised March 29, 1999

A new information matrix $[F]$ *with elements* $F_{mn} = \langle (y_m - a_m)(y_n - a_n)(\partial \ln p(\mathbf{y} \mid \mathbf{a})/\partial a_m)(\partial \ln p(\mathbf{y} \mid \mathbf{a})/\partial a_n) \rangle$ *is analyzed. The PDF* $p(\mathbf{y} \mid \mathbf{a})$ *is the usual likelihood law.* $[F]$ *differs from the Fisher information matrix by the presence of the first two factors in the given expectation. These factors make* F_{mn} *unitless, in contrast with the Fisher information. This lack of units allows* F_{mn} *values from entirely different phenomena to be compared as, for example, Shannon information values can be compared. Each element* F_{mn} *defines an error inequality analogous to the Cramer–Rao inequality. In the scalar case* $F_{mn} \equiv F$, *for a normal* $p(y \mid a)$ *law* $F = 3$, *while for an exponential law* $F = 9$. *A variational principle* $F = \min$ *(called FMIN) allows an unknown PDF* $p(x)$ *to be estimated in the presence of weak information. Under certain conditions F obeys a "Boltzmann F-theorem"* $\partial F/\partial t \leq 0$, *indicating that F is a physical entropy. Finally, the trace* $\mathcal{F}$ *of* $[F]$ *may be used as the scalar information quantity in an information-based principle for deriving distribution laws p of physics.*

1. INTRODUCTION

The concept of Fisher information[1] is a renowned and invaluable tool of estimation theory. However, Fisher information has a number of properties which are not ideal for various purposes. For example, in describing a parameter a with units (say) of length, Fisher information has units of $1/\text{length}^2$. Or, for a parameter with units of mass, the information has units of $1/\text{mass}^2$. Then a comparison of the two information values is somewhat like comparing apples with oranges. If, on the other hand, the information was unitless, such a comparison could be effectively made (as, for example,

[1] Optical Sciences, University of Arizona, Tucson, Arizona, 85721; e-mail: friedenr@super.arizona.edu.

0015-9018/99/1000-1521$16.00/0 © 1999 Plenum Publishing Corporation

when comparing Shannon information values [Eq. (48) below] from different phenomena, all of which are unitless).

The subject of this paper is, then, the development of an information quantity that is related to the Fisher information but, by contrast, has the attribute of being units-free. The new information will have other desirable properties as well.

We start out, in Section 1.1, with an enumeration of the basic properties obeyed by Fisher information, including sensitivity to units. Then in Section 1.2 we show that Fisher information does not remain invariant in form to a transformation to Fourier space. Motivated by these nonideal properties of Fisher information, we introduce and develop in Section 2 and beyond the new, unitless information measure.

1.1. Basic Measurement Theory

Suppose that we want to estimate a set of parameters $\mathbf{a} = a_1,..., a_k$ from data $\mathbf{y} = y_1,..., y_N$. The data have a random component that is usually called "noise." The estimates are chosen functions $\hat{\mathbf{a}} = \hat{a}_1(\mathbf{y}),..., \hat{a}_k(\mathbf{y})$ of the data called "estimator functions." A central issue is how accurate these estimates can be, in the presence of given noise (as specified by a PDF $p(\mathbf{y} \mid \mathbf{a})$ called the "likelihood law"). If the estimators are correct on average ("unbiased") over all possible data $\mathbf{y}$, obeying

$$\langle \hat{a}_k(\mathbf{y}) \rangle = a_k \tag{1}$$

then the mean-square errors $e_k^2 \equiv \langle (a_k - \hat{a}_k(\mathbf{y}))^2 \rangle$ of estimation obey

$$e_k^2 \geqslant [I^{-1}]_{kk} \tag{2}$$

This is the Cramer–Rao inequality.[1] Quantity $[I^{-1}]_{kk}$ is the kth element along the diagonal of the inverse matrix $[I^{-1}]$ to the Fisher information matrix $[\mathbf{I}]$, where

$$[I] \equiv \{I_{mn}\}, \qquad I_{mn} \equiv I_{mn}(p) \equiv \left\langle \frac{\partial \ln p}{\partial a_m} \frac{\partial \ln p}{\partial a_n} \right\rangle, \qquad p \equiv p(\mathbf{y} \mid \mathbf{a}) \tag{3}$$

Notation $I_{mn}(p)$ means that I_{mn} is a functional of p.

Equation (2) shows that the elements of $[I]$ must have units that are the reciprocal of those of the square of any parameter a_k. This has the undesirable effect of making it difficult to compare in some meaningful way the Fisher information values for two sets of data arising from different phenomena. By comparison, Shannon information, which is unitless, does

not present this problem. One bit of information has the same meaning regardless of phenomenon.

In physical applications mentioned below, we will be most interested in the case where there is one measurement per parameter, and the noise values x_n are independent. Then dimension $K = N$ and

$$y_n = a_n + x_n, \quad \hat{a}_n(\mathbf{y}) = \hat{a}_n(y_n), \qquad n = 1,..., N \tag{4}$$

The Cramer–Rao relationship (2), (3) now becomes more simply

$$e_n^2 I_n \geqslant 1 \tag{5}$$

$$I_n \equiv \int dy_n\, p_n(y_n \mid a_n) \left[\frac{\partial \ln p_n(y_n \mid a_n)}{\partial a_n} \right]^2 \tag{6}$$

(All integrals have infinite limits.) The tie-in with general matrix $[I^{-1}]$ preceding is that element

$$[I^{-1}]_{mn} = 1/I_n \tag{7}$$

Further, by the form of the first Eq. (4) the fluctuations in y_n purely follow those in x_n, so that

$$p_n(y_n \mid a_n) = p_{X_n}(y_n - a_n \mid a_n) \tag{8}$$

where the latter is the PDF for x_n. Assume also that p_n obeys Galilean (or, shift) invariance,

$$p_{X_n}(y_n - a_n \mid a_n) = p_{X_n}(y_n - a_n) \tag{9}$$

Using this in Eq. (6) and changing integration variable to $x = y_n - a_n$ gives

$$I_n \equiv I_n(p_{X_n}) \equiv \int dx\, \frac{(dp_{X_n}(x)/dx)^2}{p_{X_n}(x)} \tag{10}$$

With $n = 1,..., N$, this defines a vector of information quantities.

An associated *scalar* information I may be formed as the sum over all information components I_n. This is also the trace of the Fisher information matrix $[I]$. Then, by Eq. (10),

$$I = \sum_{n=1}^{N} \int dx\, \frac{(dp_{X_n}(x)/dx)^2}{p_{X_n}(x)} = 4 \sum_{n=1}^{N} \int dx \left(\frac{dq_n(x)}{dx} \right)^2 \tag{11}$$

$$\equiv I(q_1^2,..., q_N^2), \qquad \text{where} \quad p_{X_n}(x) \equiv q_n(x)^2 \tag{12}$$

defines real probability amplitude functions $q_n(x)$. We see in Eq. (11) that I no longer depends upon the parameter a_n. This is beneficial since a_n was unknown, by hypothesis.

1.2. Noninvariance of Form Under Fourier Transformation

Information form I of Eq. (11) has been used as the key element in an approach for deriving distribution laws p_{X_n} or amplitudes q_n of physics.[2–7] The approach is called the principle of extreme physical information (EPI). The trace information form I [first suggested by Stam[8]] defines the quality of data values in (initially) x-space. However, in quantum mechanics x-space and momentum μ space are equally valid specifiers of a particle, and nature is indifferent to the space chosen for the measurement. Hence, the information that is used to specify the quality of the measurement should be symmetric to the choice of space. That is, the information should have the same form in each space. However, as we show next, information I does not have this property.

Consider the case where the number N of measurements is 2. For convenience, pack the amplitude functions (q_1, q_2) as the real and imaginary parts of a complex probability amplitude

$$\psi \equiv q_1 + iq_2, \qquad i = \sqrt{-1} \tag{13}$$

Assume that the unknown PDF obeys Galilean property (9), so that I has the form Eq. (11). By the use of Eq. (13), it is simple to show that Eq. (11) may be expressed in terms of the *complex* amplitude function $\psi(x)$ as

$$I = 4 \int dx\, |\psi'(x)|^2, \qquad \psi'(x) = d\psi/dx \tag{14}$$

A complementary space to the space x is the Fourier space μ, which is specified by an amplitude function $\phi(\mu)$ obeying[3, 4, 7]

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int d\mu\, \phi(\mu)\, e^{-i\mu x/\hbar} \tag{15}$$

For a particle, Fourier space has the physical significance of being momentum space, with μ the momentum. When one substitutes Eq. (15) into Eq. (14), the result is

$$I = \frac{4}{\hbar^2} \int d\mu\, \mu^2\, |\phi(\mu)|^2 \tag{16}$$

This is not of the same form, in its dependence upon coordinate μ, as is Eq. (14) in its dependence upon coordinate x.

Hence, to attain the required invariance in form requires a different information measure.

2. A UNITLESS INFORMATION MEASURE

For simplicity, we first seek a unitless, *scalar* measure of information F. This can be independently derived in three different ways, from different aspects of the measurement process: (1) the use of unbiased data, (2) Fisher information I for logarithmic noise, and (3) Fisher information I for a net PDF $b^2x^2p_X(x)$, $b=\text{const}$. We conclude this section by constructing a unitless information *matrix* F_{mn} defined as the mean outer product of a certain vector. Its scalar version is the scalar F just mentioned.

2.1. Information F from Unbiasedness Property

This derivation follows analogous steps to the standard derivation[1] of the Cramer–Rao inequality. Start by assuming one measurement y that is unbiased,

$$0 \equiv \langle y-a \rangle \equiv \int dy(y-a)\, p, \qquad p \equiv p(y \mid a) \tag{17}$$

Differentiate this $\partial/\partial a$, giving

$$\int dy(y-a)\frac{\partial p}{\partial a} - \int dy\, p = 0 \tag{18}$$

Using normalization of p and the usual identity for the derivative of a logarithm gives

$$\int dy(y-a)\, p\frac{\partial \ln p}{\partial a} = 1 \tag{19}$$

In preparation for the use of Schwarz' inequality, we split up the integrand as

$$\int dy[\sqrt{p}]\left[(y-a)\sqrt{p}\frac{\partial \ln p}{\partial a}\right] = 1 \tag{20}$$

(This choice of splitting is the departure from the usual Fisher information derivation.) The left-hand side is the inner product of the two bracketed terms. Using the Schwarz inequality for this product gives

$$1 \leqslant \left[\int dy\, p\right]\left[\int dy (y-a)^2\, p \left(\frac{\partial \ln p}{\partial a}\right)^2\right] \tag{21}$$

Using normalization for p and the definition of an expectation, we get

$$1 \leqslant \left\langle (y-a)^2 \left(\frac{\partial \ln p}{\partial a}\right)^2 \right\rangle \equiv F(p) \equiv F \tag{22a}$$

or

$$F = \int dy (y-a)^2 \frac{1}{p}\left(\frac{\partial p}{\partial a}\right)^2 \geqslant 1 \tag{22b}$$

The latter is by use of the identity for the derivative of a logarithm. This defines the new, unitless information scalar F. In general, it is a functional of PDF p. Its unitless nature is easily verified. We note that it is always at least value unity.

At this point, it is not clear why F should be regarded as an information. It has not, e.g., been related to an error measure or a signal/noise ratio. This will be remedied in Sections 2.2–2.3.

Consider the case

$$y = a + x, \qquad p(y \mid a) = p(y-a) \tag{23}$$

of Galilean invariance for the PDF. Equation (22b) becomes

$$F(p) = \int dx\, x^2 \frac{1}{p}\left(\frac{dp}{dx}\right)^2 \geqslant 1, \qquad p \equiv p(x) \tag{24}$$

Note that this may be expressed in the alternative forms

$$F(p) = \left\langle \left(x \frac{d \ln p}{dx}\right)^2 \right\rangle \tag{25}$$

$$F(q^2) = 4 \int dx\, x^2 \left(\frac{dq}{dx}\right)^2 \tag{26}$$

Comparison of Eqs. (24) and (10) shows that the difference between the Fisher information I and information F is the factor x^2. This cancels the units $\propto 1/x^2$ of I, making F unitless as required.

Derivations (17)–(26) tacitly assumed a continuously differentiable PDF p present. For a discrete probability law $P(y_n \mid a)$ similar results follow, integrals replaced by sums. Equation (24) directly becomes

$$F = \left\langle \left[(y_n - a) \frac{\partial}{\partial a} \ln P(y_n \mid a) \right]^2 \right\rangle \tag{27}$$

[We trust that the reader will not confuse the usual notation y_n for a discrete random variable with the data notation y_n mentioned in Eq. (1) and elsewhere. The two notations are never used together.]

2.2. Examples of F for Various PDFs

As we have seen, F is a unitless number for any PDF. It is usual to attach a fictitious unit to a unitless quantity, so as to identify what kind of quantity (in this case, what type of information) it is. We propose calling this unit the "stam," in honor of a pioneer worker[8] in the field of parameter estimation.

In Table I, the values of F for some elementary scalar laws $p(y \mid a)$, a the mean, are displayed. All other parameters, such as σ in the normal law case, are arbitrary and fixed. Since the symmetric Cauchy case (with parameter b) has an intrinsic a of zero, the PDF was shifted to a finite

Table I. Information F and S/N Ratio for Various PDFs[a]

PDF law	F as $\langle \cdot \rangle$	F value	$S/N \equiv r$
Exponential	$\langle (y-a)^4 \rangle / a^4$	9 stam	1
Geometric	$\dfrac{\langle (y_n - a)^4 \rangle}{a^2(a+1)^2}$	$10 - r^2$	$\sqrt{\dfrac{a}{a+1}}$
$N(a, \sigma^2)$	$\langle (y-a)^4 \rangle / \sigma^4$	3	a/σ
Poisson	$\langle (y_n - a)^4 \rangle / a^2$	$3 + (1/r^2)$	$\sqrt{a}$
Binomial	$N^2 \dfrac{\langle (y_n - a)^4 \rangle}{a^2(N-a)^2}$	$3 - \dfrac{6}{N} + \dfrac{(N+r^2)^2}{N^2 r^2}$	$\sqrt{\dfrac{Na}{N-a}}$
$\chi^2(N)$	$N^2 \langle (y-a)^4 \rangle / 4a^4$	$3 + (6/r^2)$	$\sqrt{N/2}$
Cauchy	$\left\langle \dfrac{4(y-a)^4}{[b^2 + (y-a)^2]^2} \right\rangle$	3/2	0

[a] The mean value is a in all cases. Parameter N is used in a generic sense; N = number of trials in the binomial case or N = number of degrees of freedom in the χ^2 case.

mean value a. The F values were computed using definitions (22a) and (27), giving rise to the expectations shown in the second column. These usually involve the fourth central moment of the PDF and are evaluated in the third column. The PDFs in Table I are assumed to obey property (17) but not necessarily the Galilean property (23).

Table I shows that many PDF laws, such as the normal and exponential laws, have an F value that is a pure number, independent of the parameter(s) of the law. For other laws, information F depends upon the signal-to-noise (S/N) ratio $r \equiv a/\sigma$. This dependence is specified in the last column in Table I in terms of the mean a and other parameters of the law. This dependence upon r makes sense since r is unitless and, by its definition, F must be unitless.

In this dependence upon the S/N, F resembles the information of Shannon, which often depends (logarithmically) upon S/N. On the other hand, except for the extra factor x^2 in Eqs. (25) and (26), F would be the Fisher information. Hence, F occupies a niche somewhere between Fisher information and Shannon information. Furthermore, like the Fisher variety, F is a measure of estimation error and data error. These are shown in Sections 2.3 and 2.5, respectively.

The Fisher information I for each PDF in the table turns out to be value $1/\sigma^2$, the reciprocal of its variance, except for the shifted Cauchy case, for which $I = 1/2b^2$. Note that all these Fisher values have definite units, i.e., the reciprocal of the square of the units of y.

2.3. *F* from Fisher *I* for Logarithmic Transformation

Information F can also be shown to be the Fisher information for logarithmically transformed noise. Assume that

$$y = a + x \tag{28}$$

i.e., one shift-added data value (4). Representation (10) for I is equivalent to

$$I = \int dx\, p_X(x) \left(\frac{dp_X(x)/dx}{p_X(x)} \right)^2 \tag{29}$$

Now do a transformation of random variable from x to u, where

$$u = \exp(x/\sigma) \tag{30}$$

and σ is the standard deviation of x. Using (30) in (29) gives

$$I=\frac{1}{\sigma^2}\int du\, p_U(u)\left(1+u\frac{p'_U(u)}{p_U(u)}\right)^2 \tag{31}$$

Squaring out the integrand and integrating term by term, we find that the first term is 1 by normalization, the cross term gives -2 after an integration by parts, and the last term is $F(p_U)$ by Eq. (25). Thus, Eq. (31) becomes

$$I\equiv I(p_X)=\frac{1}{\sigma^2}(F(p_U)-1) \tag{32}$$

Since by its definition (10) $I\geqslant 0$, Eq. (32) verifies result (22) that $F\geqslant 1$. Solving (32) for F gives

$$F(p_U)=1+\sigma^2 I(p_{\sigma\ln U}) \tag{33}$$

showing that information F is linear in the Fisher information I for the transformed variable.

Fisher information is called an "information" because it defines the quality of the data as measured by the error e^2 of estimation given in Eq. (5). We may now verify that F is likewise an information. By Eqs. (5) and (32),

$$e^2\geqslant\frac{\sigma^2}{F(p_U)-1},\qquad F(p_U)\geqslant 1 \tag{34}$$

Hence, the greater F is, the smaller is the estimation error, verifying that F is an information.

2.4. *F* as a Fisher Information

Here we show that F for a PDF p is equivalent to the information I for an associated PDF x^2p. This will allow F to be used in place of I in the EPI principle described previously. It will also allow conditions to be found for which F obeys an "F-theorem" analogous to the Boltzmann H-theorem.

Given a PDF $p(x)$, define an associated PDF $\tilde{p}(x)\equiv b^2x^2p(x)$, $b=\text{const.}$ (b is needed merely for normalization purposes). From this we see that

$$\frac{(d\tilde{p}/dx)^2}{\tilde{p}}=b^2\left[\frac{x^2(dp/dx)^2}{p}+4p+4x(dp/dx)\right] \tag{35}$$

Assuming that $\tilde{p}(x)$ obeys Galilean invariance, I obeys Eq. (10). Then by Eq. (35)

$$I(\tilde{p}) = b^2 \left(F(p) + 4 + 4 \int dx\, x(dp/dx) \right) \tag{36}$$

after deleting the subscript and using normalization for $p(x)$ and the definition (25) of F. Or

$$F(p) = I(b^2 x^2 p)/b^2 \tag{37}$$

after an integration by parts of the far-right integral in Eq. (36), and use of normalization, shows that it cancels the term $+4$.

2.5. Information Matrix [F]

The definition (22a) of F is the scalar case of a more general information matrix $[F]$. Assume that the data $\mathbf{y}$ are independent, and with one measurement y_n for each parameter value a_n. This amounts to a square data-parameter problem $K = N$. As in the derivation[1] of the Fisher information matrix, define a vector of length $(N+1)$,

$$\mathbf{v} \equiv \begin{bmatrix} y_1 - a_1 \\ (y_1 - a_1)\,\partial \ln p/\partial a_1 \\ \vdots \\ (y_N - a_N)\,\partial \ln p/\partial a_N \end{bmatrix}, \qquad y_1 - a_1 \equiv \varepsilon_1 \tag{38}$$

Next, form $\langle \mathbf{v}\mathbf{v}^{\mathrm{T}} \rangle$, the expectation of the outer product of $\mathbf{v}$. This becomes

$$\left[\begin{array}{c|cccc} e_1^2 & M_{12} & 0 & \cdots & 0 \\ \hline M_{12} & & & & \\ 0 & & F_{mn} & & \\ \vdots & & & & \\ 0 & & & & \end{array}\right], \qquad e_1^2 = \langle \varepsilon_1^2 \rangle \tag{39}$$

$$F_{mn} \equiv F_{mn}(p) \equiv \left\langle (y_m - a_m)(y_n - a_n) \frac{\partial \ln p(\mathbf{y}\,|\,\mathbf{a})}{\partial a_m} \frac{\partial \ln p(\mathbf{y}\,|\,\mathbf{a})}{\partial a_n} \right\rangle$$

after some algebra (see the Appendix). The value of element M_{12} depends upon the particular law $p(\mathbf{y}\,|\,\mathbf{a})$. Particular cases are $M_{12} = 1$ for a Poisson law and $M_{12} = 2a$ for an exponential law (see the Appendix). Elements F_{mn}

in (39) define an information matrix $[F]$ which is similar in form to the Fisher information matrix (3). For our data scenario, the elements of F may be evaluated,

$$[F]=\begin{bmatrix} F_{11} & 1 & \cdot & \cdot & 1 & 1 \\ 1 & F_{22} & 1 & \cdot & \cdot & 1 \\ 1 & 1 & F_{33} & \cdot & \cdot & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & F_{KK} \end{bmatrix} \tag{40}$$

all 1's except for $diag\{F_{mn}\}$ (see the Appendix). Each element F_{mn} reverts to the scalar information $F(p)$ given by Eq. (25).

By its construction as an outer product, $\langle \mathbf{v}\mathbf{v}^{\mathrm{T}} \rangle$ is a positive-definite matrix. Therefore the determinant of matrix (39) is positive. Expanding by cofactors along the top row of (39) gives

$$e_1^2 \det[F] - M_{12}^2 Cof(F_{11}) \geqslant 0 \tag{41}$$

More generally, the approach (38) onward can be taken with a top element in vector $\mathbf{v}$ which is $y_n - a_n = \varepsilon_n$. This permits us to replace subscripts 1 by n in Eq. (41). The result is

$$e_n^2 \geqslant M_{12}^2 F_{nn}^{-1}, \qquad e_n^2 \equiv \langle \varepsilon_n^2 \rangle \tag{42}$$

The data error goes inversely with $[F]$, verifying that $[F]$ is indeed an information quantity. This inequality may be compared with the Cramer–Rao inequality (2).

3. TRANSFORMATION PROPERTIES OF F

Information F has some important invariance properties, next derived and discussed.

3.1. Invariance to Change of Scale

We continue with the independent data case of the preceding section. Suppose that a new set of units is chosen for the parameters a_n and data y_n, so that

$$y_n = C_n z_n, \qquad a_n = C_n b_n, \qquad C_n = \text{consts} \tag{43}$$

Then

$$\hat{a}_m(\mathbf{y}) - a_m \equiv y_m - a_m = C_m(z_m - b_m) \tag{44}$$

Also, by elementary Jacobian theory[9]

$$p(\mathbf{y} \mid \mathbf{a}) = p_Z(\mathbf{z} \mid \mathbf{b}) \det[J(\mathbf{z}/\mathbf{y})]$$

where

$$\det[J(\mathbf{z}/\mathbf{y})] = \prod_{n=1}^{N} (1/C_n) \tag{45}$$

is the Jacobian of the transformation (43). Equations (45) give

$$\frac{\partial \ln p(\mathbf{y} \mid \mathbf{a})}{\partial a_m} = \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_m} \frac{\partial b_m}{\partial a_m} = \frac{1}{C_m} \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_m} \tag{46}$$

by Eqs. (43). Substituting Eqs. (44) and (46) into the definition (39) of F_{mn} gives

$$\begin{aligned} F_{mn}(p) &= \left\langle C_m(z_m - b_m)\, C_n(z_n - b_n) \frac{1}{C_m} \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_m} \frac{1}{C_n} \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_n} \right\rangle \\ &= \left\langle (z_m - b_m)(z_n - b_n) \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_m} \frac{\partial \ln p_Z(\mathbf{z} \mid \mathbf{b})}{\partial b_n} \right\rangle \\ &= F_{mn}(p_Z) \end{aligned} \tag{47}$$

The information is the same in the new set of units.

It is noted that neither Fisher information I nor the Shannon entropy

$$H \equiv -\int d\mathbf{y}\, p(\mathbf{y} \mid \mathbf{a}) \ln p(\mathbf{y} \mid \mathbf{a}) \tag{48}$$

obeys this property. At the other extreme, the Shannon mutual information S and the Kullback–Leibler entropy (KL) are invariant under any change of coordinate system, even nonlinear ones.[10] However, this might be an unnecessarily strong invariance property. From a measurement standpoint, it is *sufficient* to require invariance only in transforming from x-space to Fourier (momentum) μ-space, or vice versa: the choice of measurement space is arbitrary to the observer, and measurement spaces are generally Fourier conjugate spaces. A price paid for the very strong invariance

properties obeyed by informations S and KL is apparently that they cannot be used in a variational principle for deriving the general PDF laws of physics [except for those of statistical mechanics[11]]. Informations I and F, on the other hand, can be used for this purpose, as discussed below.

3.2. Invariance to Conjugate Space

The information scalar I was formed, at Eq. (11), as the trace of the Fisher information matrix. The measurements were assumed to be independent, in a scenario of one measurement per parameter a_n, and with Galilean invariance obeyed. Continuing with this measurement scenario, we likewise form an associated information scalar $\mathscr{F}$ as the trace of the elements F_{mn} The latter are given in Eqs. (39). The result is simply a sum of information quantities (26),

$$\mathscr{F}=4\sum_{n=1}^{N}\int dx\, x^2\left(\frac{dq_n(x)}{dx}\right)^2\equiv\mathscr{F}(q_1^2,\ldots,q_N^2) \tag{49}$$

[Compare with Eq. (11) for I.]

As at Eq. (13), assume $N=2$ measurements, and pack (q_1, q_2) as the real and imaginary parts, respectively, of a complex amplitude function ψ. Then Eq. (49) directly simplifies to

$$\mathscr{F}=4\int dx\, x^2\,|\psi'(x)|^2 \tag{50}$$

[compare with Eq. (14)]. We now use the Fourier relation (15) between $\psi(x)$ and conjugate function $\phi(\mu)$ to represent $\mathscr{F}$ in conjugate space.

Differentiating Eq. (15) gives

$$\psi'(x)=-\frac{i}{\sqrt{2\pi}\,\hbar^{3/2}}\int d\mu\,\mu\phi(\mu)\,e^{i\mu x/\hbar} \tag{51}$$

Using this in Eq. (50) and interchanging orders of integration gives directly

$$\mathscr{F}=\frac{4}{2\pi\hbar^3}\int d\mu\,\mu\phi(\mu)\int d\mu'\,\mu'\phi^*(\mu')\int dx\,x^2e^{(ix/\hbar)(\mu'-\mu)} \tag{52}$$

The last integral is $-2\pi\hbar^3\delta''(\mu'-\mu)$ where $\delta(\mu)$ is the Dirac delta function. Using its sifting property in Eq. (52) gives

$$\mathscr{F}=-4\int d\mu\,\mu\phi(\phi)\frac{d^2}{d\mu^2}[\mu\phi^*(\mu)] \tag{53}$$

Integrating by parts yields

$$\mathscr{F} = 4 \int d\mu \frac{d}{d\mu}(\mu\phi) \frac{d}{d\mu}(\mu\phi^*) = 4 \int d\mu\, |\mu\phi' + \phi|^2 \tag{54}$$

after evaluating the derivatives. Squaring out gives

$$\mathscr{F}/4 = \int d\mu\, \mu^2\, |\phi'|^2 + \int d\mu\, |\phi|^2 + \int d\mu\, \mu(\phi'\phi^* + \phi'^*\phi) \tag{55}$$

We now show that the third integral cancels the second.

An integration by parts shows that

$$\int d\mu\, \phi'(\mu\phi^*) = -\int d\mu\, \phi(\mu\phi'^* + \phi^*) \tag{56}$$

Adding to this equation its complex conjugate gives

$$\int d\mu\, \mu(\phi'\phi^* + \phi'^*\phi) = -2 \int d\mu\, |\phi|^2 - \int d\mu\, \mu(\phi\phi'^* + \phi^*\phi') \tag{57}$$

The far-left and far-right hand terms are the same. This allows us to solve for it. Placing it in Eq. (55) gives the desired result,

$$\mathscr{F} = 4 \int d\mu\, \mu^2\, |\phi'|^2 \tag{58}$$

This equation and Eq. (50) are of the same form, showing that information $\mathscr{F}$ is invariant to transformation between conjugate spaces. [We previously showed, in Eqs. (14) and (16), that Fisher information I does not have this property.] Thus, measurements y made in *either* of the two spaces give rise to the same form of information $\mathscr{F}$, as required.

4. ESTIMATING PDFS USING FMIN PRINCIPLE

Consider a scenario where there are insufficient data to uniquely determine a PDF. Let the PDF obey Galilean invariance so that F obeys Eq. (26). Then F increases with the gradient content of $q(x)$. Hence, if F is minimized through choice of $q(x)$, it will produce a $q(x)$ that has minimal gradient content. Then $q(x)$ is maximally smooth and, by $p = q^2$, $p(x)$ will

be maximally smooth as well. A consequence is that $p(x)$ is minimally biased toward particular values of x. This is a desirable property for an estimated PDF (see below). How, then, may F be minimized?

As one possibility, F may be incorporated (see below) into the EPI principle[4,7] for deriving physical PDFs. That approach assumes prior knowledge of the physical source that is strong enough to imply an *exact answer* for the PDF.

A second possibility, and the one that concerns us here, is to minimize F subject to imprecise information, that is, prior knowledge that is not sufficient (as above) to produce an exact answer for the PDF. It can only be estimated, as opposed to the error-free derivation afforded by EPI. Any estimated PDF should be minimally biased toward particular x values, since such bias can often be artifactual. Hence, a smooth estimate is desired.

Let the prior information about $p(x)$ be in the form of expectations

$$d_n = \langle k_n(x) \rangle \equiv \int dx\, k_n(x)\, q^2(x), \qquad n = 1, \ldots, N \tag{59}$$

where $k_n(x)$ is a known kernel function. Knowledge of such equalities may be built into the extremization approach by use of the Lagrange method of undetermined multipliers.[12] The result is a principle,

$$\int dx\, x^2 \left(\frac{dq}{dx}\right)^2 + \sum_{n=1}^{N} \lambda_n \int dx\, k_n(x)\, q^2(x) = \min \tag{60}$$

The factor 4 in F has been absorbed into the multipliers λ_n. We call this the FMIN principle. Some particular solutions of interest are found next.

Consider the important case where

$$k_n(x) = x^{n-1} \tag{61}$$

so that the data d_n are moments. Substituting Eq. (61) into principle (60) and using the Euler–Lagrange solution gives a formal solution,

$$\frac{d}{dx}\left(\frac{\partial \mathscr{L}}{\partial q'}\right) - \frac{\partial \mathscr{L}}{\partial q} = 0 \tag{62a}$$

$$\mathscr{L} = x^2 \left(\frac{dq}{dx}\right)^2 + \sum_{n=1}^{N} \lambda_n x^{n-1} q^2 \tag{62b}$$

We carry through some particular solutions.

Suppose that $N=1$. This represents the weakest level of constraint information: a normalization constraint. The minimum obtained in F will then be the absolute one. Substituting Eq. (62b) into (62a) gives a differential equation,

$$x^2q'' + 2xq' - \lambda q = 0, \qquad \lambda \equiv \lambda_1 \tag{63}$$

The solution that allows for normalization is[13]

$$\begin{gathered} q(x) = Ax^{-\alpha}, \qquad p(x) = A^2x^{-2\alpha} \\ \alpha \geqslant 1/2, \qquad x_0 \leqslant x \leqslant x_1, \qquad \alpha, x_0, x_1 = \text{const} \end{gathered} \tag{64}$$

This is a monotonic, and therefore smooth, law as was required. Evaluating A by the normalization requirement and substituting the resulting $p(x)$ into Eq. (25) gives $F=4\alpha^2$ (independent of x_0 and x_1). Since, by (64), the minimum possible value of α is $\frac{1}{2}$, this means that the minimum possible value of F is unity [verifying Eq. (22)]. We may note that this result holds even in the limit as $x_0 \to 0$ and $x_1 \to \infty$, i.e., as the curve $q(x)$ is allowed to extend over the entire positive real line.

That the answer for $p(x)$ with $\alpha_{\min} = \frac{1}{2}$ is a simple $1/x$ law is of interest. Such a PDF is the only one that is invariant to a change of units.[14] This might have been expected since F is itself unitless, and the constraint of normalization is the weakest one possible. The $1/x$ law has previously been found[15] to describe the occurrence of unrelated numbers (weights, lengths, times, etc.) binned in a common histogram. Since the fundamental physical constants are, by definition, the most unrelated numbers of all, they might likewise be expected to follow the $1/x$ law. In fact, this is the case.[7, 14]

A $1/x$ law has been suggested[16] to be the appropriate prior probability law for defining a random variable about which nothing is known except normalization and its positivity. Steps (60)–(64) amount to a derivation of this hypothesis. To our knowledge, (60) is the only variational principle that naturally (under minimal constraints) gives rise to a $1/x$ law.

The $1/x^\alpha$ result followed from FMIN in the weakest possible constraint case. This implies that, with stronger constraints, FMIN solutions will *tend* toward $1/x^\alpha$. This is verified next.

Consider the case $N=3$, corresponding to constraints on normalization, the first moment a, and the variance σ^2. In place of Eq. (63) we now have the differential equation

$$x^2q'' + 2xq' - \lambda_1 q - \lambda_2 xq - \lambda_3 x^2 q = 0 \tag{65}$$

The normalized solution that also obeys the three constraints is the γ- (or chi-square)-law

$$p(x) = Ax^{\beta}e^{-\gamma x}, \qquad x \geqslant 0$$
$$A = \left(\frac{r^2}{a}\right)^{r^2} \Big/ \Gamma(r^2), \qquad \beta = r^2 - 1 \tag{66}$$
$$\gamma = r^2/a, \qquad r \equiv a/\sigma$$

[Note: This assumes that $a \geqslant 0$. If instead $a \leqslant 0$, then the domain is $x \leqslant 0$ and x is replaced by $(-x)$ on the right-hand side.] The γ-law answer is a unimodal, smooth function, as required. For comparison, the answer $p(x)$ given by maximum entropy H or by EPI under the same constraint information is the normal law.[4] This is a smooth law as well. Which solution to choose is arbitrary, until other considerations are made. If, for example, the user is convinced that the three constraints are *sufficient* to define the PDF, then he should choose the EPI approach.

Solution (66) is a $1/x^{\alpha}$-type law (64) modified by an exponential falloff. This shows that the bias toward $1/x^{\alpha}$ spoken of above is a real, persistent effect. Also, the dependence upon the given parameters a, σ is only through a and the signal/noise ratio r. Once again, signal/noise is a natural specifier of F-theory.

5. USE OF $\mathscr{F}$ IN EPI PRINCIPLE

The key information quantity in the EPI principle is the Fisher trace information I defined in Eq. (11). We show, next, that I may be replaced by a particular choice of the F-trace information $\mathscr{F}$. Comparing the *q-form* expressions (11) for I and (49) for $\mathscr{F}$, and using identities (26) and (37), shows that

$$I \equiv I(q_1^2, \ldots, q_N^2) = b^2 \mathscr{F}(q_1^2/b^2x^2, \ldots, q_N^2/b^2x^2) \tag{67}$$

By the use of this identity, EPI may be reexpressed in terms of the new information $\mathscr{F}$. This is beneficial because of the favorable transformation properties that $\mathscr{F}$ was previously found to have. The result is a revised EPI principle based upon an information that is independent of the choice of measurement space and of the units employed.

6. BOLTZMANN F-THEOREM

It was recently found[17] that, for temporally dependent PDFs $p(x \mid t)$, the Fisher information obeys an "I-theorem,"

$$\partial I/\partial t \leqslant 0, \qquad I = I(p) \tag{68}$$

This indicates an increase in disorder (as measured by I) with time and is analogous to the Boltzmann H-theorem. The proviso to (68) is that p obey the Fokker–Planck differential equation. Then the correspondence (37) between I and F allows us to state an "F-theorem,"

$$\partial F/\partial t \leqslant 0, \qquad F = F(p) \tag{69}$$

provided that the associated PDF $\tilde{p} \equiv x^2 p$ obeys the Fokker–Planck equation. Result (69) indicates that F is a physical measure of disorder.

7. SUMMARY

A unitless variant F on Fisher information was found to arise out of an approach (17)–(22) analogous to the usual derivation of the Cramer–Rao inequality and Fisher information. This gives the scalar (single datum) version of F. Correspondingly, an F-information *matrix* (39) or (40) arises out of a multiple-data scenario. This follows an approach (38)–(39) analogous to the usual derivation of the Fisher information matrix. Another result of interest is the error inequality (42), analogous to the Cramer–Rao inequality.

Information F is computed for various PDFs, as given in Table I. The F value for one-parameter laws are often pure numbers (e.g., $F = 3$ stam for a normal law). For multiple-parameter laws, F depends upon the S/N ratio (a unitless quantity).

Information F is related to the Fisher information I in certain direct ways. For example, F is linearly related to I [Eq. (33)] when the added noise x in Eq. (28) is expressed as the logarithm of an associated random variable. Also, F for a PDF p equals I for an associated PDF proportional to $x^2 p$ [Eq. (37)].

The information matrix $[F]$ is found [Eqs. (43)–(47)] to be invariant to a change of scale.

Information F may be used in a variational principle (60), called FMIN, for estimating an unknown PDF in the presence of insufficient information. The PDF should obey Galilean invariance. The solution to

FMIN under only the normalization constraint is a PDF (64) of the form $1/x^{2\alpha}$. The minimum obtained for F attains its absolute minimum value, of 1, for the case $\alpha = 1/2$. The resulting $1/x$ law is known to be invariant to a linear change of scale, corroborating a property previously found for F. Also this PDF has been proposed in the past to define the appropriate prior probability for a number whose only known property is that of positivity.[16] Finally, this law defines the histogram of unrelated quantities, such as masses, lengths, times, etc., and consequently describes the histogram of the fundamental physical constants.[7, 14]

The solution to FMIN under constraints of normalization and the first two moments is a gamma- or chi-square law (66). This modifies the $1/x^{2\alpha}$ behavior previously found with an exponential dropoff.

Information F obeys a Boltzmann F-theorem (69) provided that the associated PDF $x^2 p$ obeys the Fokker–Planck equation. Then F is a measure of *physical* disorder, like entropy H.

Just as the trace of the Fisher information matrix is taken to define the scalar information I, the trace of $[F]$ may be taken to define a scalar information $\mathscr{F}$ Eq. (49)]. This information has the important property that its representation in Fourier (e.g., momentum) space is the same as its representation in direct (e.g., position) space [compare Eqs. (50) and (58)]. Finally, because of Eq. (37), EPI may be expressed in terms of the information scalar $\mathscr{F}$ and, hence, in terms of an information that is of the same form regardless of the (arbitrary) choice of measurement space.

APPENDIX: CALCULATION OF SOME MATRIX ELEMENTS

The aim is to compute the elements of $\langle \mathbf{v}\mathbf{v}\rangle^{\mathrm{T}}$, with vector $\mathbf{v}$ given by Eq. (38). This is in the presence of independent and unbiased data,

$$p(\mathbf{y} \mid \mathbf{a}) = \prod_{n=1}^{N} p(y_n \mid a_n) \tag{A1a}$$

$$\langle y_n \rangle = a_n \tag{A1b}$$

Since vector $\mathbf{v}$ is of dimension $(N \times 1)$ the outer product matrix $\langle \mathbf{v}\mathbf{v}^{\mathrm{T}} \rangle$ is of dimension $N \times N$.

By sight, the (1, 1) element is $\langle (y_1 - a_1)^2 \rangle \equiv \langle \varepsilon_1^2 \rangle \equiv e_1^2$.

The (1, 2) element is called M_{12}. It is obvious that element $M_{21} = M_{12}$. Directly from (38),

$$M_{12} \equiv \left\langle (y_1 - a_1)^2 \frac{\partial \ln p(\mathbf{y} \mid \mathbf{a})}{\partial a_1} \right\rangle = \left\langle (y_1 - a_1)^2 \frac{\partial \ln p(y_1 \mid a_1)}{\partial a_1} \right\rangle \tag{A2}$$

by independence (A1). We may now drop subscript 1. The answer for M_{12} depends specifically upon the form of the law $p(y \mid a)$. For a Poisson law $p(y \mid a) = \exp(-a)\, a^y/y!$, we have $\partial \ln p/\partial a = (y-a)/a$ so that, by (A2),

$$M_{12} = \frac{\langle (y-a)^3 \rangle}{a} = 1 \tag{A3}$$

[The third central moment of the Poisson law is the mean.[18]] Or, for an exponential law $p(y \mid a) = a^{-1} \exp(-y/a)$, $y \geqslant 0$, we have $\partial \ln p/\partial a = (y-a)/a^2$; then by (A2)

$$M_{12} = \frac{\langle (y-a)^3 \rangle}{a} = 2a \tag{A4}$$

The general form for elements F_{mn} of $[F]$ as indicated in (39) directly results from the outer product of vector $\mathbf{v}$ defined in Eq. (38). we now verify that, under the independence conditions (A1a), (A1b) the elements of $[F]$ are as given in (40). First consider the off-diagonal elements F_{mn}, $m \neq n$. Expanding out the derivative of a logarithm and using independence effect (A1a), definition (39) collapses into a product of one-dimensional integrals,

$$F_{mn} = F_m F_n, \qquad F_n = \int dy_n \frac{\partial p_n}{\partial a_n} (y_n - a_n)$$

$$p_n \equiv p(y_n \mid a_n) \tag{A5}$$

We can express

$$F_n = \frac{\partial}{\partial a_n} \int dy_n\, y_n p_n - a_n \frac{\partial}{\partial a_n} \int dy_n\, p_n \tag{A6}$$

By the unbiasedness property (A1b) the first integral is a_n, and by normalization the second integral is 1. Therefore, $F_n = 1$, so that by (A5) $F_{mn} = 1$. This verifies the off-diagonal elements in (40).

Going through the same procedure for the case $m = n$, the multiple integral-mean in (39) collapses into a one-dimensional integral,

$$F_{mn} = \int dy_n\, p(y_n \mid a_n)(y_n - a_n)^2 \left(\frac{\partial \ln p(y_n \mid a_n)}{\partial a_n} \right)^2 \tag{A7}$$

This is the same as the scalar answer (22a) and, in case (23) of Galilean invariance for p_n, readily goes over into the form (25).

REFERENCES

1. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 1968).
2. B. R. Frieden and R. J. Hughes, *Phys. Rev. E* **49**, 2644–2649 (1994).
3. B. R. Frieden, in *Advances in Imaging and Electron Physics, Vol. 90*, P. W. Hawkes, ed. (Academic, Orlando, FL, 1995), pp. 123–204.
4. B. R. Frieden and B. H. Soffer, *Phys. Rev. E* **52**, 2274–2286 (1995).
5. B. R. Frieden and W. J. Cocke, *Phys. Rev. E* **54**, 257–260 (1996).
6. W. J. Cocke, *Phys. Fluids A* **8**, 1609–1614 (1996).
7. B. R. Frieden, *Physics from Fisher Information* (Cambridge University Press, Cambridge, 1998).
8. A. J. Stam, *Inf. Control* **2**, 101–112 (1959).
9. B. R. Frieden, *Probability, Statistical Optics and Data Testing*, 2nd edn. (Springer, New York, 1991).
10. F. Reza, *An Introduction to Information Theory* (McGraw–Hill, New York, 1961).
11. Ref. 4, Appendices A and C.
12. G. A. Korn and T. M. Korn, *Mathematical Handbook*, 2nd edn. (McGraw–Hill, New York, 1968).
13. E. Kamke, *Differentialgleichungen Losungsmethoden und Losungen* (Chelsea, New York, 1948).
14. B. R. Frieden, *Found. Phys.* **16**, 883-906 (1986).
15. F. Benford, *Proc. Am. Philos. Soc.* **78**, 551–572 (1938).
16. H. H. Jeffreys, *Theory of Probability*, 3rd edn. (Oxford University Press, London, 1961).
17. A. R. Plastino and A. Plastino, *Phys. Rev. E* **54**, 4423–4426 (1996).
18. Ref. 9, p. 50.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium

AFRL-SR-BL-TR-00-

0280

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 4/12/00	3. REPORT TYPE AND DATES COVERED Final Report – 9/10/98 – 10/31/99

4. TITLE AND SUBTITLE
Topics in Unconventional Imagery

5. FUNDING NUMBERS
F49620-98-1-0228

6. AUTHOR(S)
B. Roy Frieden, Principal Investigator
Sergio Barraza-Felix, Doctoral Graduate Student

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Optical Sciences Center
University of Arizona
P.O. Box 210094 - Tucson, Arizona 85721

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Airforce Office of Scientific Research/NE
110 Duncan Ave., Room B115
Bolling Airforce Base
Washington, DC 20332-8080

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

A new approach for digitally reducing the presence of random atmospheric turbulence in imagery was developed. This 'image division' method is based upon the use of two short-exposure images as data. These have unknown point spread functions (PSFs) as defined by the random turbulence. The latter are found by dividing the two image spectra (thus 'image division'), producing a set of linear equations which may be inverted for the PSFs. The inversion problem is inherently ill-posed. However, the solution is stabilized by imposing prior knowledge about the object and two PSFs : that (a) they have finite support extensions and (b) they numerically obey positivity. Once the two PSFs are found, they are used to inverse-filter their corresponding images. The result is two output reconstructions of the object, which are simply averaged to produce the final output. The approach was successfully applied to both simulations and to real infrared imagery from Kitt Peak National Observatory.

14. SUBJECT TERMS
Satellite imagery; missile imagery; imagery through turbulence; blind deconvolution; image division method; wind shear detection

15. NUMBER OF PAGES
41

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

MEMORANDUM OF TRANSMITTAL

Airforce Office of Scientific Research/NE
110 Duncan Ave., Room B115
Bolling Airforce Base
Washington, DC 20332-8080

☐ Reprint (Orig + 2 copies)

☐ Manuscript (1 copy)

☐ Technical Report (Orig + 2 copies)

☒ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: F49620-98-1-0228

REPORT TITLE: **Final Report – Topics in Unconventional Imagery**

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

B. Roy Frieden

B. Roy Frieden, Professor
The University of Arizona
1630 E. University Blvd.
Tucson, Arizona 85721

Enclosure 4